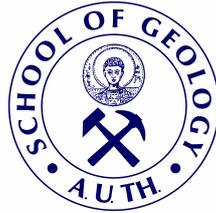




ARISTOTLE UNIVERSITY OF THESSALONIKI

SCHOOL OF GEOLOGY

DEPARTMENT OF GEOPHYSICS



POLYZOIS P. BOUNTZIS

BS Mathematics, MSc Statistics and Modeling

**Modeling of earthquake occurrence times through  
Markov Processes: A contribution to seismic hazard  
assessment in Greece**

Ph.D. THESIS

THESSALONIKI

2022





Α. ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΓΕΩΛΟΓΙΑΣ



ΠΟΛΥΖΩΗΣ Π. ΜΠΟΥΝΤΖΗΣ

Πτυχιούχος Μαθηματικός, MSc στη Στατιστική και Μοντελοποίηση

Μοντελοποίηση των χρόνων γένεσης σεισμών με  
χρήση Μαρκοβιανών διαδικασιών: Συμβολή στην  
εκτίμηση της σεισμικής επικινδυνότητας στον  
ελληνικό χώρο

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΘΕΣΣΑΛΟΝΙΚΗ

2022





POLYZOIS P. BOUNTZIS

BS Mathematics, MSc Statistics and Modeling

**Modeling of earthquake occurrence times through  
Markov Processes: A contribution to seismic hazard  
assessment in Greece**

Presented to the Department of Geophysics, School of Geology, Aristotle  
University of Thessaloniki

Submitted to the School of Geology in July 2022

Defence Date: 13/07/2022

Annex Number of the Scientific Yearbook of the School of  
Geology N: 230

**Advisory Committee**

Professor Papadimitriou Eleftheria, Supervisor

Professor Tsaklidis George, Member of the Advisory Committee

Assistant Professor Votsi Irene, Member of the Advisory Committee

**Examining Committee**

Professor Papadimitriou Eleftheria

Professor Tsaklidis George

Assistant Professor Votsi Irene

Professor Karakostas Vasileios

Assistant Professor Vasiliadis Georgios

Professor Lippiello Eugenio

Senior Researcher Zúñiga F. Ramón



All rights reserved.

Modeling of earthquake occurrence times through Markov Processes: A contribution to seismic hazard assessment in Greece – *Ph.D. Thesis*

©Πολυζώης Π. Μπουντζής, 2022

Με επιφύλαξη παντός δικαιώματος.

Μοντελοποίηση των χρόνων γένεσης σεισμών με χρήση Μαρκοβιανών διαδικασιών: Συμβολή στην εκτίμηση της σεισμικής επικινδυνότητας στον ελληνικό χώρο  
*Διδακτορική Διατριβή*

Citation:

Bountzis, P., 2022. – Modeling of earthquake occurrence times through Markov Processes: A contribution to seismic hazard assessment in Greece. Ph.D. Thesis, School of Geology, Aristotle University of Thessaloniki, Annex Number of Scientific Annals of the School of Geology No 230, 236 pp.

Μπουντζής, Π. 2022. – Μοντελοποίηση των χρόνων γένεσης σεισμών με χρήση Μαρκοβιανών διαδικασιών: Συμβολή στην εκτίμηση της σεισμικής επικινδυνότητας στον ελληνικό χώρο. Διδακτορική Διατριβή, Τμήμα Γεωλογίας Α.Π.Θ., Αριθμός Παραρτήματος Επιστημονικής Επετηρίδας Τμ. Γεωλογίας Νο 230, 236 σελ.

It is prohibited to copy, store and distribute this work, in whole or in part, for commercial purposes. Reprinting, storing and distributing for non-profit, educational or research purposes is permitted, provided the source is acknowledged and the present message retained. Questions regarding the use of the work for profit should be addressed to the author.

The approval of this Ph.D. Thesis by the school of Geology of the Aristotle University of Thessaloniki does not imply acceptance of the opinions of the author (L. 5343/1932, Article 202, par. 2)



This research is 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 project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY)



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







## Preface

The aim of this doctoral dissertation is to study the short and long-term properties of the temporal distribution of seismicity through the use of stochastic modeling. The dissertation is structured in four chapters.

In Chapter 1, we provide a short introduction that includes the motivation to the considered problem and our scientific goals. We present a description of the current state of related work in the literature concerning the Hidden Markov Models and the clustering algorithms that are applied in earthquake catalogs, and we give the theoretical background of studies related to the existence of non-stationarity for the occurrence of large earthquakes. There is also an introduction to the selected study areas which satisfy the requirements for each application. We give a description about the seismotectonic properties of the areas, as well as the seismic activity with emphasis on the strong earthquakes.

In Chapter 2, we provide the necessary mathematical background and we describe in detail the Markovian Arrival Process (MAP) model along with its analytical properties. We present a brief overview on the available estimation methods and focus on the Expectation-Maximization algorithm, which we use in the thesis. We introduce a grid-based method for the determination of the initial parameter set that we implement in a parallel-framework for reducing the required computational time. Next, we deploy the procedure for simulating data sets with the MAP model and we give the estimators of the hidden state probabilities of the model with the use of the forward-backward equations. We provide a series of simulations to validate their stability. Finally, we present the tools for model selection and for the evaluation of the process.

In Chapter 3, we introduce the idea of MAP to be used as a change point tool for seismicity rates and we investigate its efficiency on earthquake catalogs. Then, we establish a two-step clustering algorithm that combines the MAP model with a Density-Based Clustering Algorithm (DBSCAN). We believe that



incorporating a model with multiple embedded occurrence rates, the MAP, we can separate potential seismic clusters from background seismicity. We evaluate its performance on a simulated earthquake catalog where the structure of the clusters is known a-priori. Finally, we apply the method to three major seismic zones of Greece and investigate their clustering properties.

In Chapter 4, we introduce a two-step modeling procedure based on the extreme values of the observations to reveal the long-term properties of large earthquakes temporal distribution. The aim of this approach is to assess if the MAP model with non-stationary characteristics contributes to the forecasting of the large earthquakes number. We implement pseudo-prospective experiments based on simulations of the earthquake temporal distribution for the comparison of the proposed model against the Poisson, non-Poisson renewal models and the temporal Epidemic Type Aftershock Sequence (ETAS) model.

Finally, Chapter 5 concludes the thesis presenting and summarizing the research results.



## Acknowledgements

First of all, I would like to express my gratitude to my academic supervisor Prof. Eleftheria Papadimitriou for her guidance and support, as, without her, I would not have been able to complete this endeavor. Her endless energy and commitment to hard work motivated me the most. I appreciate her strong prompt for collaborating and visiting other institutions and for participating in scientific conferences. I would also like to thank my supervisor Prof. George Tsaklidis for his scientific support and advice when I most needed him during the challenges I faced. I appreciate Asst. Prof. Irene Votsi for her help and guidance during my thesis and I would also like to express my gratitude to Prof. Eugenio Lippiello from the University of Caserta, for his scientific assistance during all these years and the fruitful discussions we had.

I am also grateful to Prof. Vasileios Karakostas for his valuable contribution and collaboration regarding the seismotectonic analysis. I would like to thank my committee members Asst. Prof. Georgios Vasiliadis and Senior Researcher Zúñiga F. Ramón for evaluating my thesis and for their remarks.

A special thanks is devoted to my friends and my colleagues, Anastasios Kostoglou, Christos Kourouklas, Dr. Dimitrios Chorooglou, Dr. Ourania Mangira and Pavlos Bonatis, for being real partners all these years and for sharing with them all these efforts, thoughts and feelings. Most importantly, I am full of gratitude for my parents, my brother and my sister who have always supported me in every step of my life.



# Contents

<b>Preface</b>	<b>i</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>1</b>
<b>Περίληψη</b>	<b>4</b>
<b>Publications</b>	<b>7</b>
<b>1 Introduction</b>	<b>11</b>
1.1 Context . . . . .	11
1.2 Background and Related Work . . . . .	14
1.2.1 Markov models with hidden states . . . . .	14
1.2.2 Clustering algorithms . . . . .	18
1.2.3 Non-stationarity of large earthquakes . . . . .	25
1.3 Study areas . . . . .	28
1.3.1 Corinth Gulf area . . . . .	30
1.3.2 Central Ionian Islands area . . . . .	33
1.3.3 North Aegean Sea area . . . . .	35
<b>2 Markovian Arrival Process modeling</b>	<b>39</b>
2.1 Introduction . . . . .	39
2.2 Mathematical preliminaries . . . . .	40
2.2.1 Markov Processes . . . . .	40



2.2.2	Phase-Type distributions . . . . .	43
2.3	Markovian Arrival Process . . . . .	47
2.3.1	General case . . . . .	47
2.3.2	Special case . . . . .	51
2.3.3	Markov Modulated Poisson Process . . . . .	52
2.4	Parameter estimation . . . . .	53
2.4.1	Determination of the initial values . . . . .	58
2.5	Simulation procedure . . . . .	59
2.6	Inference of the latent states . . . . .	60
2.6.1	Stability of the local decoding algorithm . . . . .	61
2.7	Model selection . . . . .	63
2.8	Evaluation . . . . .	64
<b>3</b>	<b>Markovian Arrival Process for earthquake clustering</b>	<b>67</b>
3.1	Introduction . . . . .	67
3.2	MAP-DBSCAN method . . . . .	70
3.2.1	MAP as a tool for change point detection . . . . .	70
3.2.2	Temporal Constraints . . . . .	72
3.2.3	DBSCAN algorithm . . . . .	73
3.2.4	Evaluation of the MAP-DBSCAN method . . . . .	75
3.2.5	Generic ETAS parameters . . . . .	78
3.3	Applications . . . . .	80
3.3.1	Seismicity rate changes through MAP in Corinth Gulf, Greece . . . . .	80
3.3.1.1	Data . . . . .	81
3.3.1.2	Model evaluation and comparison with MMPP	82
3.3.1.3	Identified seismic sequences . . . . .	87
3.3.2	Application of MAP-DBSCAN . . . . .	94
3.3.2.1	Data . . . . .	96
3.3.2.2	Evaluation on synthetic catalog . . . . .	100



3.3.2.3	Potential clusters identification in three seismic zones of Greece . . . . .	106
3.3.2.4	Sensitivity analysis . . . . .	108
3.3.2.5	Cluster analysis in the three seismic zones . . . . .	113
3.3.2.6	Regional variability of clustering properties . . . . .	119
3.3.2.7	Sequence-specific clustering properties . . . . .	122
3.4	Summary and discussion . . . . .	127
3.5	Conclusions . . . . .	132
<b>4</b>	<b>Markovian Arrival Process for forecasting large earthquakes number</b>	<b>135</b>
4.1	Introduction . . . . .	135
4.2	Data . . . . .	138
4.2.1	Catalog description . . . . .	138
4.2.2	Non-stationary features of the catalog . . . . .	140
4.3	Methodology . . . . .	142
4.3.1	Establishment of MAP model with an "idle" state . . . . .	143
4.3.2	Pseudo-prospective forecasting framework . . . . .	147
4.3.2.1	Competing forecasting models . . . . .	150
4.4	Application in circum-Pacific belt . . . . .	153
4.4.1	Application of MAP . . . . .	153
4.4.2	Application of MAP with idle state . . . . .	158
4.4.2.1	Evaluation . . . . .	158
4.4.2.2	Existence of non-stationarity . . . . .	161
4.4.3	Pseudo-prospective experiments . . . . .	166
4.4.3.1	Comparison with the ETAS model . . . . .	173
4.4.3.2	Comparison with the initial MAP . . . . .	175
4.5	Discussion . . . . .	177
4.6	Conclusions . . . . .	178
<b>5</b>	<b>Concluding Remarks</b>	<b>181</b>



<b>A Appendix A</b>	<b>184</b>
A.1 ETAS formulation . . . . .	184
A.2 Estimation procedure . . . . .	187
A.3 Simulation . . . . .	188
<b>B Appendix B</b>	<b>190</b>
B.1 Gardner and Knopoff algorithm . . . . .	190
B.2 Reasenberg Linked-Based algorithm . . . . .	191
B.3 Nearest-Neighbor algorithm . . . . .	191
<b>C Appendix C</b>	<b>193</b>
C.1 Additional figures of cluster analysis . . . . .	193
C.2 Additional fitting results in circum-Pacific belt . . . . .	201
<b>References</b>	<b>202</b>





## Abstract

Strong earthquakes exhibit the largest fatality among natural hazards, posing a unique threat to the society, and causing serious damage and loss of life. In recent years an increased emphasis is given on the development of stochastic models for earthquake forecasting and the quantification of their predictive skills, which provide information that help to reveal aspects of seismogenesis and contribute to the seismic hazard assessment. This is a part of the statistical seismology research field that is focused on the statistical modeling of earthquake occurrences for the better understanding of their distribution in time, space and magnitude. The main goal of this dissertation is to propose new stochastic models and advanced statistics for the study of the short and long-term properties of seismicity in time.

Towards this direction, we introduce the use of the Markovian Arrival Process for modeling the temporal distribution of seismicity, which can be seen as a stochastic point process with intensity rate driven by a hidden Markov model. It shows large flexibility that has been emerged to be useful for capturing a large variety of behaviors and under appropriate parameterization can approximate a wide class of counting processes like the Poisson process, renewal models and more bursty ones. However, the increased flexibility of the model is linked to the large parameter set necessary for the approximation of the observed behavior sufficiently close. For the fitting of the parameters we use the Expectation-Maximization algorithm which is an appropriate approach in problems with unobserved data. We introduce a grid-based method for the choice of the initial parameter set which we implement in a parallel-framework for reducing the required computational time. One basic issue when a hidden process is applied, is to estimate the most probable sequence of latent states. For this problem, we propose the use of a local decoding algorithm that considers the forward-backward equations, and we verify its stability on simulated catalogs.



The evaluated transitions among the hidden states of the MAP model indicate changes in seismicity rate, therefore, we propose the use of the model as an off-line tool for change point detection. The identification of seismicity rate changes is important as they can be associated with seismic clusters triggered either from stress changes or fluid intrusion. We establish a two-step clustering procedure that comprises the MAP model, for an initial separation of the background seismicity from potential seismic excitations, using the revealed changes in the seismicity rate, and a density-based clustering algorithm, DBSCAN, for the detection of elevated density areas in space. We evaluate its performance on a simulated earthquake catalog where the structure of the clusters is known a-priori. Earthquake clustering is an essential aspect of short-term seismicity that can provide crucial information for the determination of faulting geometry as well as to extract useful information on the aftershock productivity of the study area and the behavior of the foreshock activity, whereas background seismicity is also essential to probabilistic seismic hazard analysis for the production of hazard maps. We detect the seismic clusters of three major seismic zones in Greece and provide their clustering properties with the use of the Epidemic Type Aftershock Sequence model, that incorporates the well established Utsu productivity law and Omori-Utsu law, respectively.

Concerning the long-term properties of seismicity, we assume that large earthquakes temporal distribution is characterized by non-stationarity, between extended periods of seismic quiescence with long inter-event times that characterize the tail of their distribution and periods of moderate seismic activity. The short-term concentration of seismicity often obscures long-term features that may characterize the earthquakes temporal distribution, therefore we consider the long inter-event times as extreme events due to their rarity and propose a two-step estimation procedure of the model, where the extreme events are estimated separately from the short-time values. Statistical analysis and forecasting in problems that incorporate extreme events is known to be highly complex as the short times do not conform well with the rare large val-



ues, and the extremes are estimated separately ignoring the potential effects of the short-time data. We provide pseudo-prospective experiments based on simulations of the earthquake temporal distribution to demonstrate the contribution of the MAP model to the forecasting of large earthquakes number and for the comparison against the Poisson, non-Poisson renewal models and the temporal ETAS model.

**Keywords:** Statistical Seismology, Stochastic Modeling, Markovian Arrival Process, DBSCAN, Seismic Clustering, Non-stationarity, Change Point Detection, Large Earthquakes, Greece, Earthquake Forecasting

## Περίληψη

Οι ισχυροί σεισμοί προκαλούν τη μεγαλύτερη θνησιμότητα μεταξύ των φυσικών καταστροφών, αποτελώντας σημαντική απειλή για την κοινωνία και προκαλώντας σοβαρές ζημιές και απώλειες ζώων. Τα τελευταία χρόνια παρατηρείται έντονο επιστημονικό ενδιαφέρον για την ανάπτυξη στοχαστικών μοντέλων με σκοπό την πρόγνωση σεισμών και την ποσοτικοποίηση των προγνωστικών τους ικανοτήτων, παρέχοντας πληροφορίες που βοηθούν στην αποκάλυψη πτυχών της σεισμογένεσης και συμβάλλουν στην εκτίμηση της σεισμικής επικινδυνότητας. Η στατιστική σεισμολογία αποτελεί ένα επιστημονικό πεδίο που επικεντρώνεται στη στατιστική μοντελοποίηση της σεισμικότητας με σκοπό την καλύτερη κατανόηση της κατανομής των σεισμών στο χρόνο και χώρο. Προς αυτή την κατεύθυνση είναι προσανατολισμένη η παρούσα διατριβή, με κύριο στόχο να προτείνει νέα στοχαστικά μοντέλα και προηγμένα στατιστικά εργαλεία για τη μελέτη των βραχυπρόθεσμων και μακροπρόθεσμων ιδιοτήτων της σεισμικότητας στο χρόνο.

Προτείνουμε τη χρήση του μοντέλου Μαρκοβιανών Διαδικασιών Αφίξεων (ΜΔΑ) για τη μοντελοποίηση της χρονικής κατανομής της σεισμικότητας. Το μοντέλο θεωρείται μια στοχαστική σημειακή διαδικασία με ρυθμό γένεσης που καθοδηγείται από ένα κρυφό Μαρκοβιανό μοντέλο. Παρουσιάζει μεγάλη ευελιξία καθώς υπό την κατάλληλη παραμετροποίηση μπορεί να προσεγγίσει μια ευρεία κατηγορία στοχαστικών διαδικασιών, όπως η διαδικασία Poisson, μοντέλα ανανέωσης και πιο εκρηκτικές διαδικασίες. Ωστόσο, το τίμημα για την αυξημένη ευελιξία του μοντέλου συνδέεται με το μεγάλο πλήθος παραμέτρων που χρειάζεται να εκτιμηθούν. Για την προσαρμογή των παραμέτρων χρησιμοποιήθηκε ο αλγόριθμος Expectation-Maximization που θεωρείται κατάλληλη προσέγγιση σε προβλήματα με μη παρατηρήσιμα δεδομένα. Για την επιλογή των αρχικών παραμέτρων του αλγορίθμου εισάγαμε μία μέθοδο που βασίζεται στην κατασκευή πλέγματος, την οποία υλοποιήσαμε με παράλληλο προγραμματισμό για τη μείωση του απαιτούμενου υπολογιστικού χρόνου. Ένα βασικό πρόβλημα σε διαδικασίες με κρυφές καταστάσεις είναι η εκτίμηση της πιο πιθανής ακολουθίας των κρυφών

καταστάσεων. Προτείνουμε έναν αλγόριθμο αποκωδικοποίησης με τη χρήση των εμπρός-πίσω εξισώσεων, και επαληθεύσαμε τη σταθερότητά του σε προσομοιωμένους καταλόγους.

Οι εκτιμώμενες μεταβάσεις μεταξύ των κρυφών καταστάσεων του μοντέλου ΜΔΑ υποδεικνύουν αλλαγές στο ρυθμό σεισμικότητας, επομένως, προτείνουμε τη χρήση του μοντέλου ως εργαλείο για την ανίχνευση σημείων αλλαγής στη σεισμικότητα. Ο προσδιορισμός των αλλαγών του ρυθμού σεισμικότητας είναι σημαντικός, καθώς μπορεί να συσχετισθεί με σεισμικές συστάδες που προκαλούνται είτε από μεταβολές στο πεδίο των τάσεων είτε από τη διάχυση ρευστών. Αναπτύσσουμε μία διαδικασία συσταδοποίησης δύο βημάτων που περιλαμβάνει το μοντέλο ΜΔΑ, για τον αρχικό διαχωρισμό της σεισμικότητας υποβάθρου από πιθανές σεισμικές διεγέρσεις, χρησιμοποιώντας τις εκτιμώμενες αλλαγές στον ρυθμό σεισμικότητας και έναν αλγόριθμο συσταδοποίησης με κριτήριο την πυκνότητα των δεδομένων στο χώρο, DBSCAN, για την ανίχνευση περιοχών με αυξημένη συγκέντρωση σεισμών. Αξιολογούμε την απόδοσή του σε έναν προσομοιωμένο κατάλογο σεισμών όπου η δομή των συστάδων είναι γνωστή εκ των προτέρων. Η ομαδοποίηση σεισμών αποτελεί κομμάτι της βραχυπρόθεσμης σεισμικότητας και παρέχει κρίσιμες πληροφορίες για τον προσδιορισμό της γεωμετρίας ρηγμάτων καθώς και για την παραγωγικότητα των μετασεισμικών ακολουθιών μίας περιοχής μελέτης όπως επίσης για την ύπαρξη προσεισμικής δραστηριότητας. Παράλληλα, η σεισμικότητα υποβάθρου είναι απαραίτητη για την πιθανολογική ανάλυση του σεισμικού κινδύνου. Εφαρμόζουμε τον αλγόριθμο συσταδοποίησης σε τρεις κύριες σεισμικές ζώνες του ελληνικού χώρου και υπολογίζουμε τις ιδιότητες των ανιχνευμένων συστάδων με τη χρήση του μοντέλου ETAS, το οποίο ενσωματώνει εμπειρικούς νόμους όπως τον νόμο παραγωγικότητας του Utsu και τον νόμο Omori-Utsu.

Σχετικά με τις μακροπρόθεσμες ιδιότητες της σεισμικότητας, υποθέτουμε ότι η χρονική κατανομή των ισχυρών σεισμών χαρακτηρίζεται από μη στασιμότητα, συγκεκριμένα εκτεταμένοι περιόδοι σεισμικής ηρεμίας εναλλάσσονται με περιόδους μέτριας σεισμικής δραστηριότητας. Δείχνουμε ότι η χρονική συμπεριφορά



τους δεν μπορεί να προσεγγισθεί ικανοποιητικά από το μοντέλο ΜΔΑ λόγω της παρουσίας βραχυπρόθεσμης σεισμικότητας. Η βραχυπρόθεσμη συγκέντρωση της σεισμικότητας συχνά εμποδίζει τη μελέτη μακροπρόθεσμων ιδιοτήτων που μπορεί να χαρακτηρίζουν τη χρονική κατανομή των σεισμών, επομένως θεωρούμε τους μεγάλους χρόνους μεταξύ συμβάντων ως ακραία γεγονότα λόγω της σπανιότητάς τους και προτείνουμε μια διαδικασία εκτίμησης δύο βημάτων για το μοντέλο, όπου τα ακραία φαινόμενα εκτιμώνται χωριστά από τα υπόλοιπα δεδομένα. Η στατιστική ανάλυση και πρόβλεψη σε προβλήματα που ενσωματώνουν ακραία φαινόμενα είναι γνωστό ότι είναι πολύ περίπλοκη, καθώς οι σύντομοι χρόνοι δεν συμμορφώνονται καλά με τις σπάνιες μεγάλες τιμές, και σε πολλές περιπτώσεις λαμβάνονται υπόψιν χωριστά ώστε να αποφύγουμε τις πιθανές επιπτώσεις των βραχυχρόνιων δεδομένων. Παρέχουμε προγνωστικά πειράματα βασισμένα σε προσομοιώσεις της χρονικής κατανομής των σεισμικών συμβάντων για να δείξουμε τη συμβολή του μοντέλου ΜΔΑ στην πρόβλεψη του πλήθους ισχυρών σεισμών και για τη σύγκριση με τα μοντέλα Poisson, μη Poisson μοντέλα ανανέωσης και το χρονικό μοντέλο ETAS.



## Publications

The scientific findings of this thesis have been published in peer-reviewed journals and international conference proceedings. Below we present the complete list.

### Refereed International Journals:

1. Bountzis, P., Papadimitriou, E., & Tsaklidis, G. (2020). Earthquake clusters identification through a Markovian Arrival Process (MAP): Application in Corinth Gulf (Greece). *Physica A: Statistical Mechanics and its Applications*, 545, 123655;
2. Bountzis, P., Papadimitriou, E. and Tsaklidis, G. (2021) Markovian Arrival Processes for earthquake clustering analysis. In: *Statistical Methods and Modeling of Seismogenesis* (eds. Limnios, N., Papadimitriou, E. and Tsaklidis, G.), Wiley-ISTE: London, UK, 241-270;
3. Bountzis, P., Papadimitriou, E., & Tsaklidis, G. (2022). Identification and Temporal Characteristics of Earthquake Clusters in Selected Areas in Greece. *Applied Sciences*, 12, 1908;
4. Bountzis, P., Tsaklidis, G., & Papadimitriou, E. (2022). Pseudo-prospective forecasting of large earthquakes full distribution in circum-Pacific belt incorporating non-stationary modeling. *Physica A: Statistical Mechanics and its Applications*. **Accepted.**

### Refereed Conference Proceedings:

1. Bountzis, P., Papadimitriou, E. & Tsaklidis, G., (2018). Markovian Arrival Process modeling for the detection of seismicity rate changes for the strong earthquakes in Greece. In: *Proceedings of the 31st Panhellenic Statistics Conference*, 278-291.

### Other presentations:



1. Bountzis P., Papadimitriou E. & Tsaklidis G. (2018). Detection of seismicity rate changes for the strong earthquakes of Greece by Markovian arrival process modeling In 31st Panhellenic Statistics Conference, Thessaloniki, Greece;
2. Bountzis P., Papadimitriou E. & Tsaklidis G. (2018). Markovian Arrival Process modeling of seismicity patterns in Corinth Gulf, Greece. In 36th General Assembly of the ESC, ESC2018, Valletta, Malta;
3. Bountzis P., Papadimitriou E. & Tsaklidis G. (2019). Earthquake clusters in Corinth Gulf, Greece: Identification through a Markovian arrival process (MAP). In 15th International Congress of the Geological Society of Greece, Athens, Greece;
4. Bountzis P. & Papadimitriou E. (2019). Are large earthquakes triggered on a global scale?, 32nd Panhellenic Statistics Conference, Ioannina, Greece;
5. Bountzis P. & Lippiello E. (2019). An approximation method for a fast evaluation of the Epidemic Type Aftershock Sequence (ETAS) occurrence probability. In 11th International Workshop on Statistical Seismology, Hakone, Japan;
6. Bountzis P., Papadimitriou E. & Tsaklidis G. (2019). A Phase-Type approximation of the earthquake size distribution. In 11th International Workshop on Statistical Seismology, Hakone, Japan;
7. Bountzis, P., Papadimitriou, E. & Tsaklidis G. (2021). Pseudo-prospective forecasting of large earthquakes in circum-Pacific belt incorporating non-stationary modelling. In Virtual 37th General Assembly of the ESC, ESC2021.

Peer-reviewed papers outside the scope of the thesis:



1. Bountzis, P., Papadimitriou, E. & Tsaklidis, G. (2019). Estimating the earthquake occurrence rates of Corinth Gulf through Markovian Arrival Process modeling. *Journal of Applied Statistics*, 46, 995-1020
2. Kostoglou, A., Karakostas, V., Bountzis, P., & Papadimitriou, E. (2020). The February-March 2019 Seismic Swarm Offshore North Lefkada Island, Greece: Microseismicity Analysis and Geodynamic Implications. *Applied Sciences*, 10, 4491
3. Bountzis, P., Kostoglou, A., Papadimitriou, E., & Karakostas, V. (2021). Identification of spatiotemporal seismicity clusters in central Ionian Islands (Greece). *Physics of the Earth and Planetary Interiors*, 312, 106675
4. Papadimitriou, E., Bonatis, P., Bountzis, P., Kostoglou, A., Kourouklas, Ch. & Karakostas, V. (2022). The intense 2020–2021 earthquake swarm in Corinth gulf: faulting architecture and cluster analysis from high resolution microseismicity. *Pure and Applied Geophysics*. **Under minor revision.**





# Chapter 1

## Introduction

### 1.1 Context

The need for a resilient society motivates the scientific community to intensify its efforts for reliable earthquake forecasts and advanced ground-motion models, two key ingredients for seismic hazard assessment. The development of modern methods, which will be based on established or new models, as well as new tools for the evaluation of their effectiveness and the quantification of their uncertainties, could provide a higher level of preparedness for the upcoming destructive earthquakes ([Jordan et al., 2011](#)).

Towards this direction we need to fully understand the mechanisms that generate earthquakes, a process called seismogenesis. We know that an earthquake occurs when the accumulated stress on a fault exceeds its strength, so a physics-based model that incorporates this information could predict future large earthquakes. However, the fault geometry is complex, as they form populations and secondary structures within a region which could be linked with each other ([Ben-Zion and Sammis, 2003](#)). Additionally, the physical processes responsible for the generation of large earthquakes are not directly observable and require near-fault seismic and geodetic dense sensor networks in conjunction with improved analysis techniques that can resolve multi scale processes

(Kato and Ben-Zion, 2021).

The field of statistical seismology endeavors to fill in the gap between physics-based models without statistics, and statistical modeling. The use of stochastic modeling is based on the fact that some features of the earthquake generation process remain unknown, and they are replaced within the model by a random component. It includes the quantification of the earthquakes distribution in time, space and magnitude along with their corresponding uncertainties. This information helps to reveal aspects of seismogenesis and provides probability estimates for the genesis of strong earthquakes that could contribute to the seismic hazard assessment. In a sense, stochastic modeling can be considered physical in terms of its aim to describe the physical process of earthquakes generation (Vere-Jones, 2010).

Earthquake clustering is one of the main aspects of seismicity, and is expressed by the concentration of earthquakes in time and space. Earthquakes generation is neither characterized by periodicity nor is it random over time, but exhibits strong short-term clustering and weak long-term variations (Kagan and Jackson, 1991). The first is expressed in the form of either triggered aftershocks after the occurrence of large events or earthquake swarms, whereas the latter is associated with the temporal distribution of main shocks. Clustering algorithms provide significant information on the identification of active faulting structures and the well-revealed spatiotemporal clustering of seismicity can be embedded in stochastic models to construct time-dependent earthquake forecasts (Field et al., 2017). The Collaboratory for the Study of Earthquake Predictability (CSEP) highlighted the significant probability gain of clustering models, such as the Reasenbergs–Jones model (Reasenbergs and Jones, 1994), the Short-Term Earthquake Probability model (Gerstenberger et al., 2005), and the Epidemic-Type Aftershock Sequence model (Ogata, 1998), during aftershock sequences over time-independent models (Schorlemmer et al., 2018). The knowledge on the spatiotemporal aftershock patterns is important because large aftershocks can pose their own hazard and can be used for operational



aftershock forecasting. Clustering algorithms can also be implemented to produce "declustered" earthquake catalogs, i.e., to remove all triggered events, which are subsequently given as input for long-term probabilistic seismic hazard analysis (Gerstenberger et al., 2020).

The worldwide expansion of the seismic networks along with the constant development of methods for earthquake monitoring increased the earthquake detectability and larger data sets which include smaller earthquakes, e.g. magnitude completeness  $m_c \approx 0.3$  (Ross et al., 2019) with 1.8 million events for Southern California earthquake catalog during 2008-2017, are now available. This avalanche of data can provide us with additional information regarding the clustering features of a region (foreshock activity, aftershock duration and productivity, existence of swarms) which can be used for constructing more accurate forecasts and reveal secondary faulting structures. However, it requires the development of refined and robust statistical tools for the identification of the short-term seismic clusters.

A second type of clustering concerns the long-term variations of seismicity attributed to large main shocks. Combining the intense space-time concentration of events and the short duration of the available instrumental earthquake catalogs, this type of clustering is often indistinct. The effectiveness of statistical methods for the evaluation of the large earthquakes long-term behavior might be limited in regional fault systems, and CSEP suggests to explore new forecasting models in a global scale (Schorlemmer et al., 2018). The recent surge of great earthquakes in circum-Pacific belt (Beroza, 2012; Lay, 2015), with the 2004  $M9.0$  Sumatra, the 2010  $M8.8$  Maule, Chile, and the 2011  $M9.1$  Tohoku, Japan earthquakes that constitute the half of the six largest earthquakes on record, implies the need to investigate extensively whether or not non-stationarity in large earthquakes exist.

The aim of the present PhD thesis is to study the short and long-term properties of the temporal distribution of seismicity with stochastic modeling. This thesis addresses the following research topics and proposes novel algorithms,

leveraging advanced statistics and stochastic modeling theory, that can be applied within the scope of seismicity clustering and earthquake forecasting:

- The development of a change point procedure for detecting seismicity rate changes through stochastic modeling.
- The establishment of a two-step procedure for the identification of earthquake clusters.
- The development of a two-step modeling procedure for the large earthquakes temporal distribution that incorporates non-stationary characteristics.
- The contribution of the procedure to the earthquakes number forecasting.

## 1.2 Background and Related Work

### 1.2.1 Markov models with hidden states

The Hidden Markov models (HMMs), constitute a general category of time-dependent stochastic processes with non-observed states in the data ([Baum and Petrie, 1966](#)). In many cases, the underlying earthquake mechanisms are not evaluated with an adequate accuracy, which leads to lack of knowledge regarding the causal relationship with the observed seismicity. The hidden factor embedded in a process can serve as an appropriate tool for modeling the process of seismogenesis.

New techniques and analytical tools have been developed by [Votsi et al. \(2013\)](#), who identified the unobserved stress level controlling the strong earthquakes occurrence with magnitudes  $M \geq 6.5$ , in Greece and its surrounding areas, since 1845. A discrete time Hidden semi-Markov model (HSMM) was applied for the first time by the same authors [Votsi et al. \(2014\)](#), by providing a statistical estimator of the intensity function, further extended by [Pertsinidou](#)

[et al. \(2017\)](#) who assumed different emission and jump times of the HSMM. The seismicity rates in the area of Ionian Sea, Greece, for the period 1900-2006 were estimated through a Poisson HMM (PHMM) by [Orfanogiannaki et al. \(2010\)](#). The model revealed changes of seismicity and recognized earthquake clusters with the aim to estimate future seismicity rates. Migration of seismic activity to adjacent areas within North Aegean Sea is revealed by a multivariate PHMM, which was developed by [Orfanogiannaki and Karlis \(2018\)](#). [Wu \(2010\)](#) proposed a simple HMM for earthquake declustering and compared it with the ETAS model using data from central and western Japan and developed another class of models, called quasi-HMMs, to estimate the location of the next aftershock ([Wu, 2011](#)). The HMMs have been also used for the short-term forecasting of  $M \geq 4.0$  earthquakes in Southern California region under the RELM experiment ([Ebel et al., 2007](#)) and for main shock seismic activity in southern California and western Nevada ([Chambers et al., 2012](#)). [Yip et al. \(2018\)](#) developed a version of a HMM where forecasts of the occurrence times and magnitudes of earthquakes are generated simultaneously.

The HMMs have also been used for the modeling of volcanic eruptions. [Bebbington \(2007\)](#) found that during a period of 406 years (1600-2006) the volcano of Mount Etna is characterized by long Poisson periods alternating frequently with periods of triggered eruptions. Tremor (non-volcanic) activity is another type of data that has been investigated with the use of HMMs. [Wang et al. \(2017\)](#) developed a type of hidden Markov models, where each state represents a distinct segment of tremor sources and revealed the existence of migration patterns of tremors within the Tokai region. They further proceeded to their classification based on the observed occurrence patterns: episodic, weak concentration, and background ([Wang et al., 2018](#)).

However, for the application of the HMMs a fixed-length time interval needs to be chosen, which can lead to missing information due to the overdispersion of earthquake occurrences. The dynamics of the generation process are often characterized by episodic trends and heterogeneity, which makes ex-

tremely demanding the correct choice for the length of the time step. Figure 1.1a illustrates an example of a data set that is characterized by fluctuations in the seismicity rate. The peaks (vertical black arrows) in the distribution of the inter-event times (Figure 1.1b), show small values with high frequency compared to their neighbors and provide visual proof on the existence of distinct periods of relatively high seismic activity. The framework of a hidden Markov process in the continuous domain where each state can be associated to each of those peaks through a distinct seismicity rate, is expected to enable the modeling of changes in the earthquake dynamics without considering a fixed-length time-step.

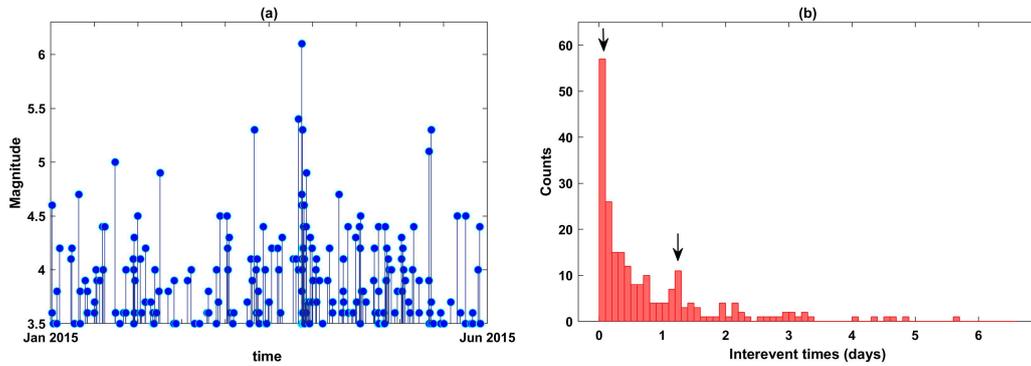


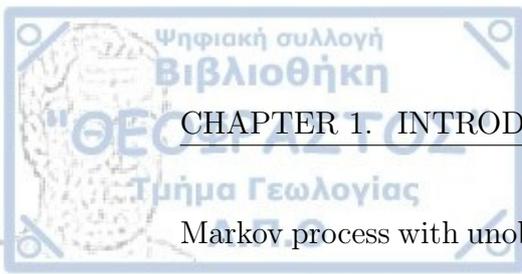
Figure 1.1: Earthquakes with  $M \geq 3.5$  that occurred in the area of Greece from January 2015 to June 2015. (a) Magnitude distribution as a function of time; (b) distribution of the inter-event times during this period. Black arrows indicate inter-event times with high frequency compared to the neighboring bins. Their values are relatively small, 0.05 and 1.25 days, respectively, so they can correspond to different periods of relatively high seismic activity.

Wang et al. (2012) proposed a Markov-modulated Hawkes process with step-wise intensity function at arrival instants for the modelling of the variations in the seismicity rate. This is a class of a self-exciting stochastic point process whose intensity function is driven by a Markov process, where hidden states correspond to distinct background and decay rates. They investigated the temporal patterns of Landers-Hector Mine series of earthquakes, where they captured the evolution of main shocks, their aftershocks duration and periods of seismic quiescence. More recently, Wang et al. (2020) pro-

posed a marked point process with varying intensity function to model the non-stationary features of the volcanic eruptions caused by missing data. A Poisson process seems insufficient for forecasting eruptions, so point processes incorporating multiple change points might be more appropriate.

The Markov Modulated Poisson Process (MMPP) [Fischer and Meier-Hellstern \(1993\)](#) is another class of time-varying intensity rate process, where each hidden state is associated with a homogeneous Poisson occurrence rate. Temporal variabilities of deep earthquake occurrences in New Zealand are investigated through MMPP [Lu and Vere-Jones \(2011\)](#) revealing the existence of active and quiescence seismicity periods. The extension of the process by adding state-dependent marks enabled the association of the occurrence times and magnitudes of New Zealand deep earthquakes with two levels of seismicity [Lu \(2012\)](#). The model has been also used to detect changes in the magnitude–frequency distribution for both deep ([Lu, 2017a](#)) and shallow ([Lu, 2017b](#)) earthquakes in New Zealand. In both cases, the variability of the b-value is found into two alternate episodes, one of relatively low and one of high b-value. In addition, the b-value is decreasing before the occurrence of large earthquakes ( $M \geq 6$ ) and periods of low b-values are mainly associated with both shallow and deep major events. Recently, the concept of a switched Poisson process for modeling the temporal features of the background seismicity in Northeastern Italy has been used by [Benali et al. \(2020\)](#), suggesting the existence of heterogeneity in the occurrence patterns. However, on a MMPP the rate of the intensity function is varying inside an arrival epoch, so that after a  $(k - 1)$ -th earthquake occurrence the latent process can move several times to different hidden states, following the  $k$ -th event with a transition to the same state. From a seismological point of view, this means that seismicity rate can only change between the occurrence of the last event and the following one, without any apparent causal relationship.

[Bountzis et al. \(2019\)](#) proposed another special case of a counting process whose occurrence rate changes step-wise at arrival instants based on a latent



Markov process with unobserved states, called Markovian Arrival Process. Under this MAP formulation the time between  $(k-1)$ -th and  $k$ -th events depends on the state of Markov process at the  $(k-1)$ -th arrival, namely, changes of the seismicity rate in MAP can only be triggered by a previous event. The inter-event times follow a hyper-exponential distribution, which is a mixture of  $n$  exponential distributions for some  $n$ . It has very tractable properties for both analytical and simulation purposes (Feldmann and Whitt, 1998) and it can capture the variability in the arrival rate, such as the one that exists in the earthquake occurrences. A mixture of exponential distributions has been adopted by Mendoza-Rosas et al. (2009) in order to provide an assessment of the volcanic hazard between explosive eruptions of the Colima and Popocatepetl volcanoes, in Mexico. Each component corresponds to a different eruption rate, characterizing in this way efficiently successive regimes of non-stationary processes and long-tail distributions. The first definition of the MAP model was given by Neuts (1979), which is considered a seminal work that extended the Markov models beyond the Exponential distribution and the Poisson process. Lucantoni et al. (1990) simplified the notation of the MAP model and Lucantoni (1991) established a general version of the current model, the Batch Markovian Arrival Process, where groups of events are allowed to occur at the same time. The MAP is a non-renewal process which due to its versatility includes as special cases the MMPP, PH-renewal process and the Poisson process among others. Cordeiro and Kharoufeh (2011) presented a wide range of MAP applications from queuing systems and inventory control theory up to telecommunications networks where the MAP models approximate the bursty behavior of the data traffic.

### 1.2.2 Clustering algorithms

Seismicity clustering incorporates many different forms of triggering mechanisms such as static and dynamic stress changes, fluid migration and aseismic

slip. The detected clustered seismicity can provide information on many different aspects of the earthquake dynamics, such as in [Lippiello et al. \(2012\)](#) who extracted the foreshocks and aftershocks of the Southern California earthquake catalog based on space-time windows and used their spatio-temporal organization to improve the forecasting of future earthquakes through a modified ETAS model. [Dascher-Cousineau et al. \(2020\)](#) related variations of the aftershock productivity to potential physical factors such as earthquake depth, lithosphere age and plate boundary type and [Shebalin et al. \(2020\)](#) showed that the productivity distribution in  $\Delta M$ , is independent of the magnitude of triggering events and decreases with depth. The Bath's law properties are extended in space, time and focal mechanisms by [Tahir et al. \(2012\)](#) who observed variations on the size and distance of the largest aftershock from the main shock with faulting type. In addition to the necessity of cluster identification, robust algorithms for earthquake declustering, i.e., the separation of the background seismicity from clustered events, contribute to the development of long-term seismic hazard maps ([Mizrahi et al., 2021](#); [Petersen et al., 2018](#); [Taroni and Akinici, 2021](#)) or the regional optimization of background rates ([Llenos and Michael, 2020](#)). Finally, some studies use seismicity clustering techniques for the reconstruction of the faulting network ([Grund et al., 2016](#); [Kamer et al., 2020](#); [Petersen et al., 2021](#)).

Seismicity clustering is quantified through the dependencies of earthquake occurrences in time, space and magnitude and is based on the existence of physical interactions such as stress changes induced by previous events ([Stein, 1999](#)) or fluid diffusion over a region ([Hainzl, 2004](#)). In the first case, it is translated into main shock–aftershock sequences, whereas in the latter case it is translated into earthquake swarms, i.e. events close in time and space without a distinct main shock. The first study of the clustering properties in time, namely, an increase of seismicity rate in a certain area, was provided by Omori in 1894 ([Omori, 1894](#)), who proposed a power-law distribution for the determination of the aftershocks number after a main shock. One century later,

Utsu *et al.* (1995) modified the functional form leading to the Omori–Utsu law. The empirical law has been incorporated into point process models, the most known one being the ETAS model (Ogata, 1988, 1998). According to the Collaboratory for the Study of Earthquake Predictability Schorlemmer *et al.* (2018) clustering models demonstrate reliable forecasts of future earthquakes, such as the ETAS model Lombardi and Marzocchi (2010), which in many cases perform much better than long-term smoothed seismicity models mainly during the aftershock sequences Omi *et al.* (2015); Page *et al.* (2016); Rhoades *et al.* (2018). However, swarms are not driven by the same physical mechanisms, which makes it difficult to extract reliable forecasting results when swarm activity is dominant in a region.

Traditional approaches for clustering detection or declustering include window-based methods where space-time windows around main shocks are defined and events within them are extracted (Felzer and Brodsky, 2006; Gardner and Knopoff, 1974; Gentili and Bressan, 2008). Peresan and Gentili (2020) studied the efficiency of the Gardner-Knopoff (GK) algorithm to separate clustered from background seismicity on earthquake catalogs from north-eastern Italy and showed that for moderate and small main shocks where aftershock activity might be low the GK method can lead to overestimated results. Teng and Baker (2019) showed that the method also provides dubious results on induced seismicity, in particular, it removed 80% of earthquakes in the Oklahoma–Kansas region and failed to approximate the observed changes in background rates.

Another approach to identify clusters of events is by creating links among earthquakes based on spatial and temporal zones (Frohlich and Davis, 1990; Reasenber, 1985). The Reasenber (RB) linked-based model (Reasenber, 1985), which is one of the most commonly used approaches especially for seismic hazard studies, assumes a spatial zone based on stress redistribution near the main shock and a temporal zone based on the Omori law. The cluster is defined by the linked events, where the largest earthquake is considered the main

shock and the others are divided into foreshocks and aftershocks. Probabilistic seismic hazard analyses are based on the RB approach for the declustering of earthquake catalogs in New Zealand (Stirling et al., 2002) and Italy (Taroni and Akinci, 2021) among others and Schorlemmer and Gerstenberger (2007) provided a range of parameter values for the RB model, that should be used for long-term forecasting models in the RELM testing center.

The Nearest-Neighbor (NN) is an approach for clustering detection which has been initially proposed by Baiesi and Paczuski (2004). They developed a space-time-magnitude metric based on empirical laws for earthquake triggering. The metric is a decreasing function of the time and space proximity among two earthquakes, i.e., the metric value for events close in time and space is smaller than the value for two events that are in long distance. The methodology has been further explored by Zaliapin et al. (2008), who introduced a rescaled distance,  $R$ , and time,  $T$ , formula among events and observed that the seismicity is divided into a clustered and a background component when  $(R, T)$  is plotted. The first component corresponds to the clustered seismicity with short inter-event values in time and space and the second component comprises the background seismicity. The key factor in their approach is the threshold value  $\eta_0$ , according to which the two components are divided. A robust analysis of the method is made to the 1981–2011 relocated seismicity catalog of southern California by Zaliapin and Ben-Zion (2013a), who demonstrated the efficiency of the algorithm. They further classified the detected clusters into burst-like and swarm-like sequences and showed the existence of correlation among the spatial variability of the clusters and the heat flow of the corresponding areas (Zaliapin and Ben-Zion, 2013b). The same authors (Zaliapin and Ben-Zion, 2016) associated the triggering mechanisms of clustering on a global scale mainly with the heat flow level, whereas the tectonic regime seems to have an indiscernible contribution (Stallone and Marzocchi, 2019). The method has been used to investigate the clustering properties of seismicity in the Sea of Marmara region, NW Turkey (Martínez-Garzón et al.,

(2019), where regions related to repeaters are identified and to Northeastern Italy (Peresan and Gentili, 2018) where two areas with distinct cluster characteristics are revealed. In particular, swarm-like sequences are associated to the north-western part and burst-like sequences with the south-eastern part of the study area. The method is also applied to swarm-type clusters by Zhang and Shearer (2016), who managed to recognize and associate swarms to fluid and heat flow as well as aseismic slip. However, uncertainties on the choice of the binary threshold led to the incorporation of a probabilistic frame by Bayliss et al. (2019), specifically the Markov Chain Monte Carlo mixture modeling, for the classification of the nearest neighbor events. Recently, Aden-Antoniów et al. (2022) suggested the use of machine learning techniques for the choice of the threshold value and the parameter tuning of the algorithm. Except of cluster detection the NN algorithm can be used for the construction of a declustering catalog (two closely related problems). Zaliapin and Ben-Zion (2020) introduced a modified version of the NN algorithm that includes the stochastic thinning for separating background and clustered events, according to a threshold value which is determined by randomized-reshuffled catalogs.

A stochastic approach for the discrimination of clustered seismicity from the background one was first implemented by Zhuang et al. (2002). The method is based on the assumption that seismicity is well described by the ETAS model and the probability for every event being a triggered or a background one is estimated through an iterative method according to the intensity function of the model. According to these probabilities, the clusters are separated from the background events and significant features of the triggered earthquakes are analyzed and revealed in various regions (Zhuang et al., 2005, 2019, 2004). Important information can be extracted also from the declustered catalog, for mapping the background seismicity or the moment rate in a seismic area like in Console et al. (2010) for the Southern California and Davoudi et al. (2018) for Iran earthquake catalogs, respectively. However, Hainzl et al. (2013) demonstrated that the  $a$ -parameter of the stationary ETAS model, which de-

termines the number of aftershocks in relation to the main shock magnitude, is underestimated in catalogs with swarm activity, while the estimation of the background seismicity component,  $\mu$ , is influenced by the existence of slow slip events (Kumazawa and Ogata, 2014). Hence, the fitting of the ETAS parameters can be biased due to transient aseismic forces and can lead to dubious cluster detection in regions that include seismic swarms. A comparison among the stochastic and the nearest-neighbor method from Varini et al. (2020) showed that they produce similar partitions of the Northeastern Italy and Western Slovenia earthquake catalogs, but they differ in the internal connections among the grouped events. The clusters derived from the stochastic method exhibit more complicated structures than the ones from the nearest-neighbor method. Another probabilistic method for separating main shocks from aftershocks is proposed by Marsan and Lengline (2008). Their procedure does not require the adoption of any particular intensity function and is non-parametric. The main assumption concerns the linearity of earthquake triggering and enables the discrimination between directly and indirectly triggered aftershocks.

There are also clustering methods that are based on the assumption of a common physical trigger during the sequence, expressed by fluctuations in the occurrence rate, such as the CURATE algorithm (Jacobs et al., 2013). The method is based principally on the comparison among the observed cumulative rate and the average one, and subsequently on the implementation of a distance and day rule without any physical constraint. The method is useful for the detection of swarm-type clustering (Mesimeri et al., 2019), where the seismic activity cannot be explicitly explained by earthquake interactions, and for identifying temporal variations of tremor activity (Peng et al., 2019). However, the choice of the free parameters should be based on the tectonic characteristics and the earthquake epicentral catalog accuracy of the region rather than being defined blindly, since it lacks a rigorous optimization procedure. Bottiglieri et al. (2009) proposes the use of the coefficient variation among inter-event

times as a tool to determine the starting and ending point of a sequence. Other statistical methods for the detection of changes in the seismicity rate concern the application of Change point Analysis (CptA). [Pievatolo and Rotondi \(2008\)](#) applied a CptA on an earthquake catalog with independent inter-event times through a Bayesian approach, whereas [Gupta and Baker \(2017\)](#) and [Fiedler et al. \(2018\)](#) developed a methodology on the spatiotemporal domain for the detection of rate changes, which they tested in induced seismicity. Recently, [Lykou et al. \(2020\)](#) applied a CptA based on non-parametric tests in the Gulf of Corinth area (Greece), with remarkably well online performance.

Finally, some studies implement two-stage clustering approaches, where they categorize separately the seismicity in terms of their temporal and spatial proximity. [Georgoulas et al. \(2013\)](#) developed a hybrid method that incorporates a density-based clustering algorithm, the DBSCAN ([Ester et al., 1996](#)), for the grouping of events concentrated in time and an agglomerative hierarchical procedure for separating the events also in space. The DBSCAN algorithm is a versatile tool for grouping events as the number and the form of the clusters is not determined a-priori. Additionally, earthquakes can be grouped according to different attributes based on the definition of the distance metric. For instance, it has been used for classifying focal mechanisms using the Kagan angle ([Cesca et al., 2014](#); [Kostoglou et al., 2020](#)), and earthquake repeaters using the waveform similarity as distance metric ([Kostoglou et al., 2021](#); [Petersen et al., 2021](#)), respectively. Another two-step clustering procedure has been proposed recently by [Vijay and Nanda \(2021\)](#). In the first step a Gaussian kernel-based temporal density estimation is used for grouping the events in time and in the second step events are also separated in space based on a decision graph of the events spatial density. For a more comprehensive review and technical details on clustering algorithms in seismology we refer to [Molchan and Dmitrieva \(1992\)](#) and [van Stiphout et al. \(2012\)](#).

### 1.2.3. Non-stationarity of large earthquakes

Kagan and Jackson (1991) in their milestone work concluded that earthquake clustering is characterized by (1) a short-term, strong clustering related to foreshock-main shock-aftershock sequences, and (2) a long-term, weak clustering of main shocks. However they noted that "investigating long-term properties of seismicity is much more difficult than similar studies of short- and intermediate-term variations of earthquake occurrence rates. One of the reasons is obvious: we lack well documented, uniform long duration catalogues of earthquakes."

Nevertheless, during the last 15 years the interest on the assumption of non-stationarity for the occurrence of large earthquakes has raised due to the noticeable increase of the great earthquakes ( $M \geq 8.0$ ) since 2004 (Beroza, 2012; Lay, 2015). Relative results indicate the existence of temporal clustering, time-dependency of great earthquakes or even the non-stationarity of the stochastic process that describes the procedure of seismogenesis, which could contribute to the improvement of these earthquakes forecasting. Bufe and Perkins (2005) displayed the existence of temporal clustering in a global scale during 1950-1965 for earthquakes with  $M \geq 8.6$ , followed by a long lasting relative seismic quiescence until 2001. They found statistically significant deviations from a stationary Poisson process through Monte Carlo simulations. The non-stationarity of earthquakes with  $M \geq 7.0$  during 1900-1990 was investigated by Lombardi and Marzocchi (2007), who found both short-term triggering activity and long-term fluctuations of the earthquake rate in decades or longer, based on the application of an ETAS model. However, the non-stationary ETAS model does not detect the times of the long-term seismicity rate changes which should be defined before. Similar results are derived by Faenza et al. (2003, 2008) who evidenced temporal clustering of large earthquakes with  $M \geq 7.0$  during 1900-2004 in a worldwide scale and with  $M \geq 5.5$  since 1600 in a regional scale (Italy), respectively. They suggest that

the physical processes that generate aftershocks and large earthquakes might be different (Faenza et al., 2004). More recently, Luginbuhl et al. (2018) concluded that global seismicity with earthquakes of  $M \geq 7.0$  from 2004 to 2016 is not random in time, after comparison of synthetic random catalogs with the observational one, through the concept of natural time, whereas Rogerson (2018) provided statistical evidence for large ( $M \geq 7.0$ ) earthquakes interactions within months between the triple junction of the Nazca, Cocos and Pacific plates. Evidence for long-range correlations among strong earthquakes ( $M \geq 6.5$ ) during 1845-2017 has also been found in the Greek region by Iliopoulos et al. (2020). They suggest the existence of interactions between strong earthquakes in intermediate time scales and in long spatial ranges.

The physical mechanisms behind the large earthquakes temporal behavior, namely, the long-term non-stationarity of the seismicity remain vague, although there has been increasing evidence for the existence of remotely triggered small earthquakes from the passage of seismic waves after a large earthquake occurrence (Hill and Prejean, 2007). Gombert and Sherrod (2014) showed that all examined earthquakes with  $M \geq 8.6$  in subduction zones since 1960 triggered  $M > 5.5$  crustal earthquakes within days and distances up to a few multiples of the dimensions of the triggering events in accordance with Parsons and Velasco (2011). However, Pollitz et al. (2012) evidenced dynamic triggering even in far field distances. Sammis and Smith (2013) suggest that these remote events usually small in magnitude could advance or delay a seismic cycle producing temporal clusters of events or in other words synchronization and Scholz (2010) observed the phase locking of nearby faults which he called fuzzy synchronization.

Bendick and Bilham (2017) statistically quantified the worldwide synchronization of earthquakes with  $M \geq 7.0$  since 1900 through topological networks and time series analyses and more recently, Bendick and Mencin (2020) revealed the existence of temporal earthquake clustering based on the renewal interval alignment by testing whether events with similar renewal intervals

tend to occur closer in time than it would be expected for independent events. Nevertheless, the Topological Data Analysis cannot be used for forecasting the productivity of large earthquakes. [Chen et al. \(2020\)](#) tried to explain the observed burstiness (clusters separated by long irregular periods of quiescence) of worldwide events with  $M \geq 6.0$  during 1904-2016 through the Devil's Staircase. It is a fractal property of non-linear complex systems where one rupture could affect the behavior in the whole area.

Many statistical studies in global earthquake catalogs reach to contradictory conclusions, i.e., earthquakes originate from a stationary Poisson process, implying that they occur randomly in time with a stable seismicity rate. [Michael \(2011\)](#) performed statistical tests to investigate whether a null hypothesis for the occurrence of random events with a constant rate along with localized aftershock sequences can be rejected for the inter-event times of  $M \geq 7.0$ , 7.5 and 8.5 earthquakes since 1900. The identified temporal clustering was attributed to the aftershock activity. Similarly, [Shearer and Stark \(2012\)](#) concluded that events originate from a Poisson process, however, they focused on the existence of correlated events in global distances neglecting potential regional-scale clustering. [Daub et al. \(2012\)](#), [Parsons and Geist \(2012\)](#) and [Ben-Naim et al. \(2013\)](#) compared the occurrence frequencies of large ( $7.0 \leq M \leq 8.3$ ) global events since 1900 with numerical simulations drawn from Poisson processes without finding any significant departure from the stationary Poisson process. [Touati et al. \(2016\)](#) applied a change point analysis for the detection of changes in the global rate of events with  $M \geq 7.0$ , 7.5, 8.0 and 8.5 since 1918, concluding that there is not any strong evidence that global earthquakes are correlated in time.

The brevity of the seismological record (duration of instrumental earthquake catalogs is very short), along with the limited number of global great earthquakes reduce the robustness of the statistical tests to identify non randomness ([Dimer de Oliveira, 2012](#)). Therefore, the seismic activity might not be representative of the potential long-term variations. [Daub et al. \(2015\)](#)

produced magnitude and time-dependent simulated catalogs and concluded that the ability of the statistical tests to distinguish random fluctuations in time from temporal clustering, depends on the type of cluster and the amount of data. Since the detection of non-stationarity in the seismicity rate might not be always possible, [Zaliapin and Kreemer \(2017\)](#) proposed the investigation of significant seismic moment temporal variations. They concluded that time-dependent parameters of the tapered Pareto distribution are in favor over time-independent models for events with  $M \geq 7.0$  during 1918-2014 and suggested that the increased global rate after December 2004 is a statistically significant feature. However, even though they robustly verified the existence of sub-intervals in time with different parameters, the duration and the number of the time-windows are rather subjectively chosen. One extra difficulty when investigating the temporal distribution of earthquakes is the co-existence of two time-scales. The short-term clustering which is expressed through foreshocks and aftershocks and the long-term variations in the seismicity rate. The strong space-time concentration in short-times might obscure the long-term features of earthquakes ([Zaliapin and Ben-Zion, 2022](#)).

### 1.3 Study areas

Greece is characterized by a complex seismotectonic environment that is dominated by the intense crustal deformation due to the subduction of the eastern Mediterranean oceanic lithosphere beneath the overriding Aegean plate (Figure 1.2), forming the Greek subduction zone and the back arc area ([McKenzie, 1972](#); [Papazachos and Comninakis, 1971](#)). The Aegean microplate accommodates a southwestward movement relative to the stable Eurasia imposing a widespread extension in the interior of the plate. A characteristic case of extension is the Corinth Gulf area (purple box in Figure 1.2).

The second major seismotectonic structure in Greek area is associated with the lateral extrusion of the Anatolian microplate as a result of its collision

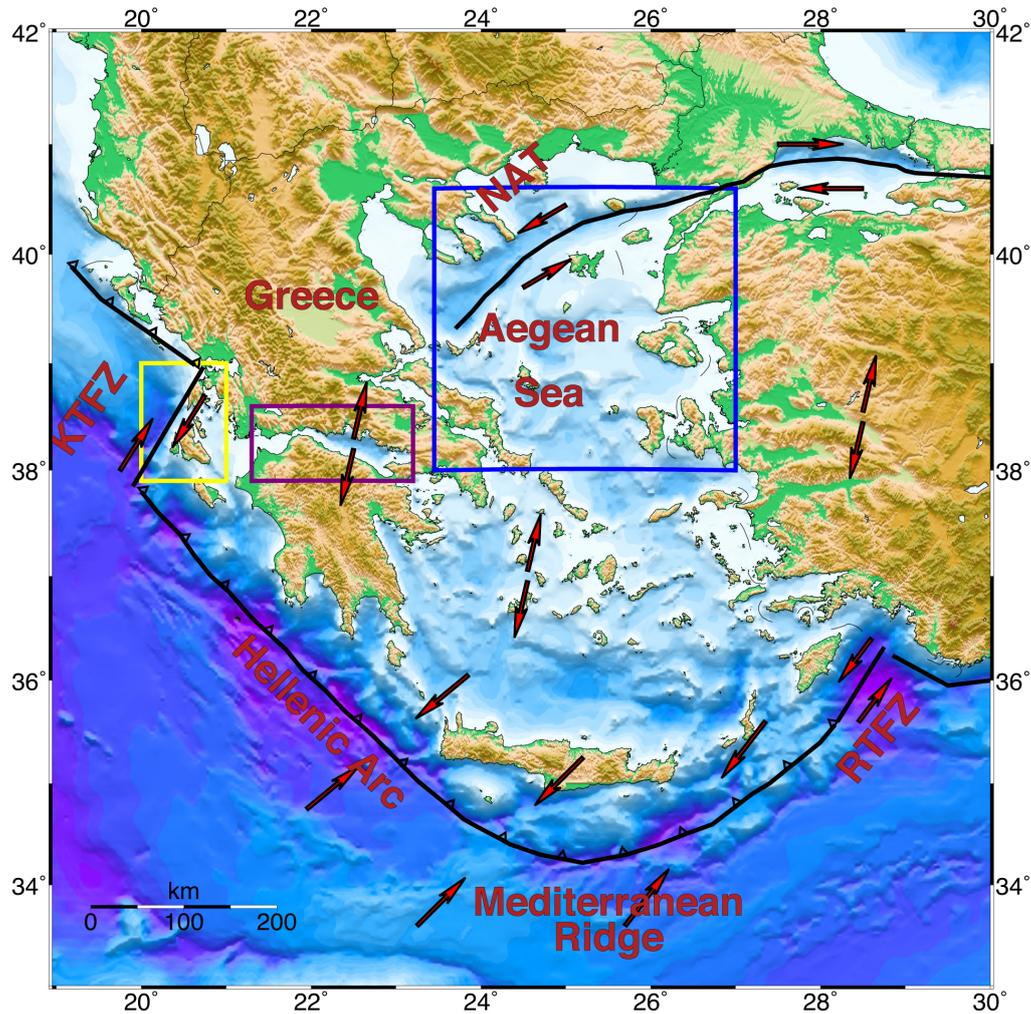


Figure 1.2: Seismotectonic properties of the Aegean Sea and the surrounding areas including the dominant structures (thick black lines), such as the Hellenic Arc (thick black line with triangles), the North Aegean Trough, which accommodates the westward prolongation of NAF into the Aegean, the Kefalonia dextral Transform Fault Zone and the Rhoades sinistral Transform Fault Zone. Red arrows represent the kinematics along major fault zones as well as the main deformational pattern in the Aegean, characterized by compression along the Hellenic Arc and extension in the back-arc area. Corinth Gulf area is confined by the purple box, the Central Ionian Islands area by the yellow box and the North Aegean Sea area by the blue box, respectively.

against the Arabian plate and the westward prolongation of the North Anatolia Fault (NAF) into the Aegean Sea (Jackson and McKenzie, 1988; Taymaz et al., 1991). The NAF is propagating to the west reaching the Aegean Sea and forming the Northern Aegean Trough (NAT) (McKenzie, 1972) (blue box

in Figure 1.2). As the right-lateral slip on the North Anatolian Fault enters the Aegean region, it distributes on several parallel faults. Dextral strike slip faulting prevails along the North Aegean Sea and its boundaries, the Kefalonia Transform Fault (KTFZ) with a dextral strike slip motion (Scordilis et al., 1985) to the west and the Rhoades Transform Fault (RTF) with a sinistral strike slip motion to the east (Papazachos and Papazachou, 2003). The two transform fault zones delimit the Hellenic Arc at its northwestern and southeastern ends, respectively. The KTFZ which can be distinguished in its two branches Kefalonia and Lefkada (Louvari et al., 1999), connects the subduction to the south and the continental collision on the west coast of Greece in the north and Albania (yellow box in Figure 1.2).

All the earthquake data sets that are used in the thesis are from the regional catalog of the Geophysics Department of the Aristotle University of Thessaloniki (Aristotle University of Thessaloniki, 1981), compiled with the recordings of the Hellenic Unified Seismological Network (HUSN) (University of Athens, 2008).

### 1.3.1 Corinth Gulf area

The Corinth Gulf (CG), constitutes one of the fastest extending continental regions in the world. The width of the gulf is not constant, but increases from west to east. Based on geodetic measurements, the rate of expansion between the western and eastern parts of the Corinth Gulf are different (Briole et al., 2000; Chousianitis et al., 2015). In particular, its western part expands with a rate of 13–14 mm/yr, whereas the eastern part exhibits a slower extension rate of 10–12 mm/yr.

Several destructive earthquakes ( $M \geq 6.0$ ) both in historical and instrumental era are observed, which verify the intense crustal deformation of the region (Ambraseys and Jackson, 1997; Papazachos and Papazachou, 2003). The seismicity is mainly associated with eight major faults that bound the rift to the

south and dip to the north (Hatzfeld et al., 2000). The Offshore Akrata, Xylokastro, Offshore Perachora, Skinos and Alepochori fault segments attribute to the eastern subarea, whereas the western one includes the Psathopyrgos, Aigion and Eliki fault segments (Figure 1.3). In addition to Corinth gulf the study area encompasses secondary structures such as Trichonis Lake, Achaias and Kapareli faults. The seismicity in Lake Trichonis graben is studied due to its proximity to the Corinth Gulf and the recent seismicity which consists of sparse activity interrupted by the occurrence of several strong earthquakes, like the 2007 earthquake swarm with  $M_w = 5.2$  (Kiratzi et al., 2008). The last strong earthquake in the study area took place in the northwestern Peloponnese on June 8, 2008 with  $M_w = 6.4$  (Ganas et al., 2009; Karakostas et al., 2017).

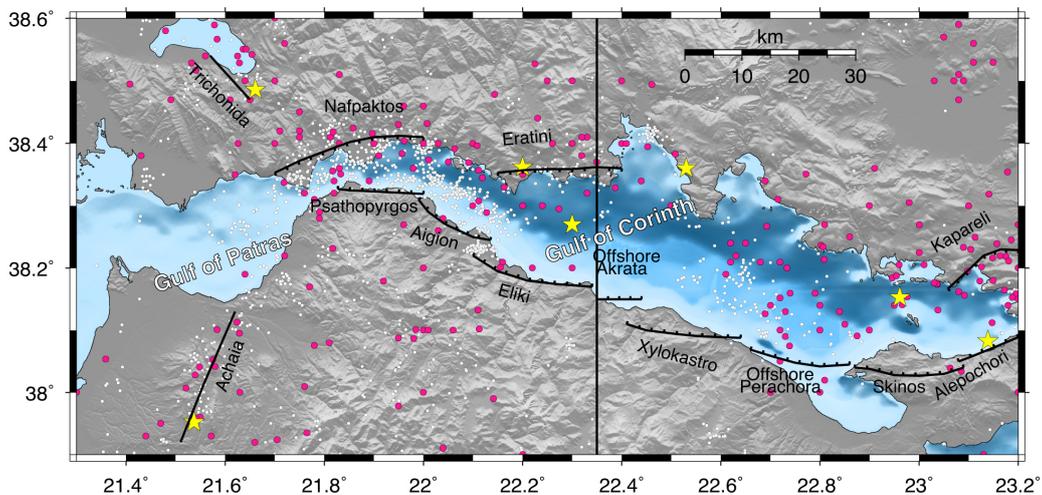


Figure 1.3: Study area with the eight major faults bounding the southern coastline. In addition, the Trichonida, Kapareli and Achaia faults edging the Corinth Gulf are given, which are associated with the 1975,  $M_w = 6.0$ , 1981,  $M_w = 6.3$  and 2008,  $M_w = 6.4$  events, respectively, and the Nafpaktos and Eratini faults. Epicentral distribution of earthquakes with  $2.5 \leq M < 4.5$  that occurred in the study area during 2012 – 2019 are shown with white circles, with  $4.5 \leq M < 6.0$  during 1964 – 2019 with pink circles and with  $M \geq 6.0$  during 1964 – 2019 with yellow stars. Vertical line divides the area into the western and eastern subareas.

The sequence of the three destructive earthquakes ( $M \geq 6.3$ ) in less than 10 days in 1981 which are associated with adjacent and antithetic faults in the eastern part of the Corinth Gulf, in Alkyonides Bay, has aroused interest and is well studied (Hatzfeld et al., 2000; Jackson et al., 1982). The western part has also been hit by destructive earthquakes such as the one near Galaxidi in 1992 (Hatzfeld et al., 1996) and in Aigio in 1995 (Bernard et al., 1997). Two moderate earthquakes with  $M_w = 5.4$  and  $M_w = 5.4$  took place near Efpalio in January 2010. They occurred within four days and at a distance of about 5km. They correspond to two adjacent faults, which probably ruptured at the same period (Ganas et al., 2013; Karakostas et al., 2012; Sokos et al., 2012). Until the end of 2019 two earthquakes with  $M \geq 5.0$  have occurred in the study area, one in November 2014 with  $M_w = 5.0$  (Kaviris et al., 2018) and one in March 2019 with  $M5.1$ .

The Corinth Gulf area is also characterized by seismic activity triggered by fluid diffusion (Bourouis and Cornet, 2009; Mesimeri et al., 2019). One of the major sequences is the 2013 Aigion swarm which initiated on 21 May 2013 with a bulk of small events and several bursts associated to earthquakes with magnitudes ranging between 3.3-3.7 (Kapetanidis et al., 2015; Mesimeri et al., 2016). Two seismic excitations followed, the first one associated to the  $M = 3.7$  event on 15 July, 2013 (Michas et al., 2021). The second half of 2014 is also a well-studied period with intense seismic activity such as the  $M4.6$  event on 21 September 2014, associated with the earthquake swarm located between Nafpaktos and Psathopyrgos (Kapetanidis, 2017).

In general, the Corinth Gulf, especially the western part, is characterized by strongly clustered seismicity that triggered the interest of many independent studies that investigated the properties of the microseismicity in the area (Duverger et al., 2018; Kapetanidis et al., 2021; Mesimeri et al., 2018a; Michas et al., 2021; Rigo et al., 1996). Finally, there is evidence for the existence of aseismic slip expressed through multiple repeating events (De Barros et al., 2020; Mesimeri and Karakostas, 2018).

### 1.3.2. Central Ionian Islands area

The central Ionian Islands (CII) constitute the most seismically active area of the Mediterranean region. Historical information and instrumental recordings evidence intense seismic activity (Papazachos and Papazachou, 2003) with strong earthquakes ( $M \geq 6.0$ ) occurring frequently, in many cases clustered in time possibly due to stress transfer and triggering of adjacent, optimally oriented fault segments (Papadimitriou, 2002). The dominant seismotectonic characteristic is the Kefalonia Transform Fault Zone extending more than 100 km along the western coastlines of Lefkada and Kefalonia Islands and comprising two distinct main branches, the Lefkada and Kefalonia fault segments. It manifests right lateral strike slip motion with a minor thrust component (Kiratzi and Langston, 1991; Scordilis et al., 1985) with NNE-SSW strike for the Lefkada and NE-SW for the Kefalonia segment (Louvari et al., 1999) (Figure 1.4). The KTFZ designates the transition between the termination of the Hellenic subduction zone to the south and the continental collision between the Adriatic and Aegean microplates to the north, causing a compressional tectonic regime with the maximum stress component at a NE-SW direction, as confirmed by both seismological (Papadimitriou, 1993) and geodetic data (Kahle et al., 1995). Microseismicity is mainly concentrated along the KTFZ and manifests both swarm-type and main shock-aftershock sequences. The investigation of the clustering features is crucial to shed light on the secondary structures and the complex geodynamics of the area. It may constitute as well a tool for exploring the evolution of the seismic sequences in the region and the preparatory phase of main shocks, contributing to the seismic hazard assessment.

The area of the central Ionian Islands offers a unique opportunity to study the clustering properties of seismicity, since the recent strong ( $M \geq 6.0$ ) earthquakes have motivated the installation of a constant monitoring of the region with a dense seismological network. The 2003,  $M_w = 6.2$ , Lefkada main shock

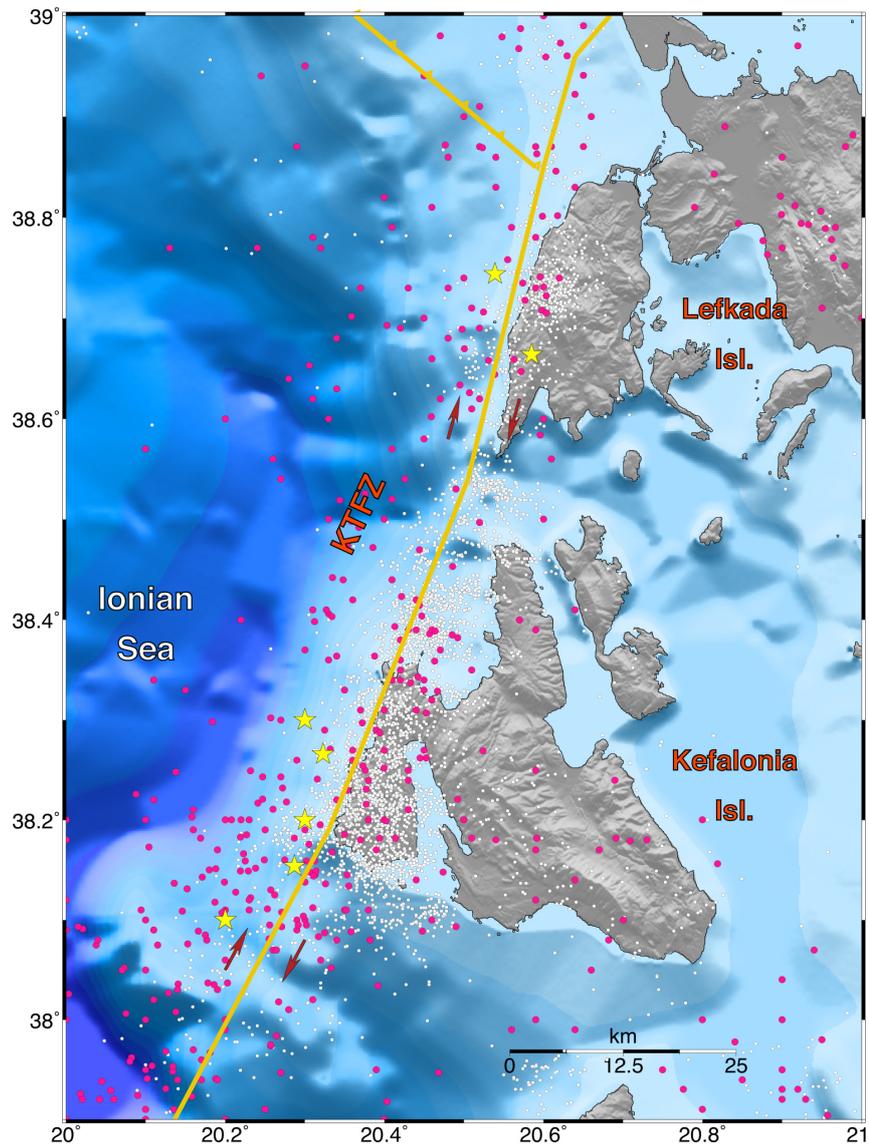


Figure 1.4: Map of the central Ionian Islands area showing the main seismic-tectonic characteristics. The thick yellow line illustrates the fault trace of the KTFZ and red arrows describe the right lateral motion in the zone. Epicentral distribution of earthquakes with  $2.5 \leq M < 4.5$  that occurred in the study area during 2012 – 2019 are shown with white circles, with  $4.5 \leq M < 6.0$  during 1964 – 2019 with pink circles and with  $M \geq 6.0$  during 1964 – 2019 with yellow stars.

and its aftershock sequence stimulated a detailed investigation of the activated fault network by an adequately accurate earthquake catalog based on the recordings of a dense portable digital network (Karakostas and Papadim-

itriou, 2010).

Among the major seismic sequences is the 2014 Kefalonia doublet with the first main shock occurring on 26 January with  $M = 6.1$ , and aftershock activity extending over 35 km, including the second main shock with  $M = 6.0$  that occurred on 3 February along with its own aftershocks. Earthquake relocation allowed for a thorough investigation of the aftershock distribution based on focal distribution and offered additional insight in the details of the rupture kinematics of the transform fault zone (Karakostas et al., 2015; Sokos et al., 2015). Another major event is the 2015  $M_w = 6.5$  Lefkada earthquake (Papadimitriou et al., 2017) 12 years after the 2003  $M_w = 6.2$  main shock that struck the northwestern part of Lefkada Island (Papadopoulos et al., 2003). The main shock ruptured a segment of the KTFZ which is adjacent to that of the 2003. The area between Lefkada and Kefalonia, extended to about 15 km, is considered as a transition zone with E–W-oriented, parallel step-over faults (Karakostas et al., 2015) that were activated in the three ruptures and are also related with smaller clusters of microseismicity (Bountzlis et al., 2021). Finally, some seismic swarm activity is observed in the northernmost terminus of the KTFZ, offshore Lefkada island (Kostoglou et al., 2020).

### 1.3.3 North Aegean Sea area

The third area is located in the North Aegean Sea (NAS) (Figure 1.5) and is dominated by dextral strike-slip faulting, along the North Aegean Trough and its parallel branches (Papazachos et al., 1998), as a consequence of the westward propagation of the North Anatolian Fault into the Aegean (McKenzie, 1972). The driving mechanism of the active deformation in the Aegean region is the subduction of the oceanic lithosphere of the Eastern Mediterranean under the continental Aegean microplate, forming the Hellenic Subduction Zone. The almost NS-oriented backward extension of the Aegean, due to the roll-back of the submerged lithospheric plate, is the driving force for the high rate

of expansion of the backward area in the Aegean (Le Pichon and Angelier, 1979). The combination of the NS expansion with the westward movement of the Anatolian plate results in rapid deformation in the wider Aegean region with NE-SW direction. The area is characterized by the frequent generation of strong earthquakes ( $M \geq 6.0$ ).

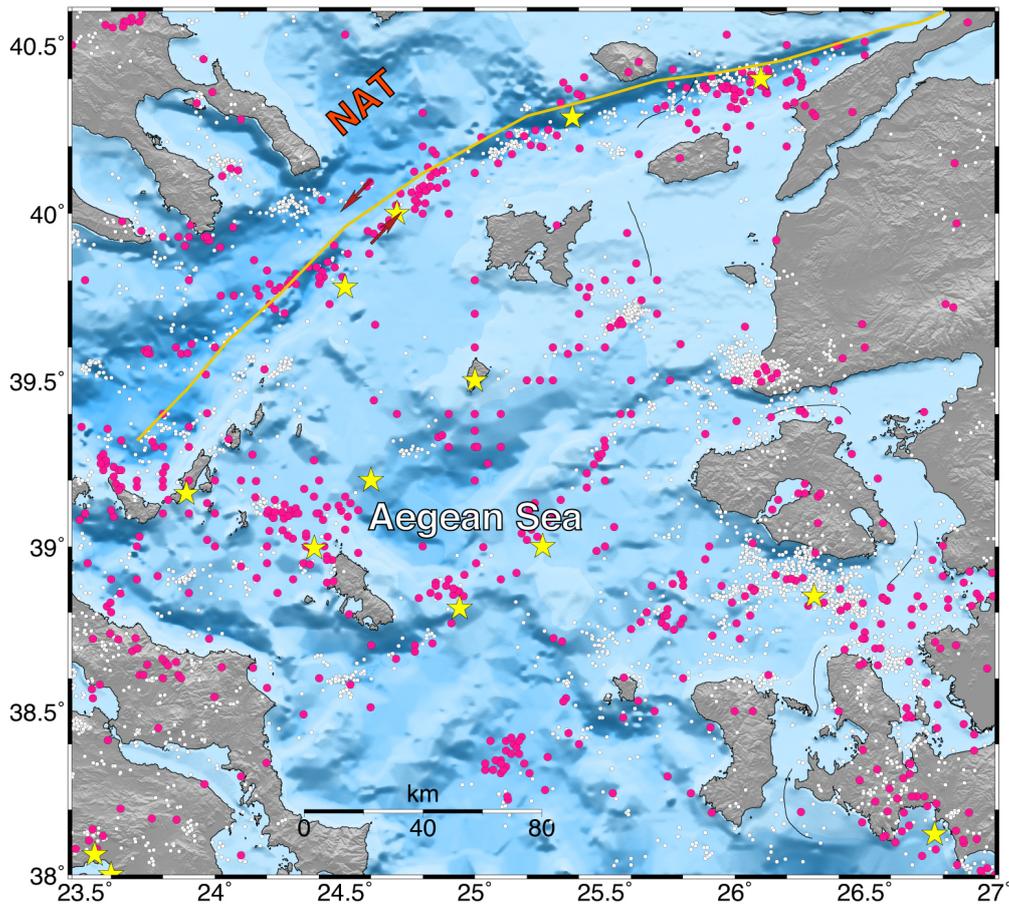


Figure 1.5: Map of the North Aegean Sea area showing the main seismotectonic characteristics. The thick yellow line illustrates the fault trace of the North Aegean Trough, which accommodates the westward prolongation of NAF into the Aegean. Epicentral distribution of earthquakes with  $2.5 \leq M < 4.5$  that occurred in the study area during 2012 – 2019 are shown with white circles, with  $4.5 \leq M < 6.0$  during 1964 – 2019 with pink circles and with  $M \geq 6.0$  during 1964 – 2019 with yellow stars.

One of the two major earthquakes in the last decade is the  $M_w = 6.9$  earthquake of May 24, 2014 that is located approximately 20 km southeast

of Samothraki island in the NAT. Its aftershock sequence is lacking of strong aftershocks with  $M > 5.0$ , as well as aftershocks very close to the main shock (Evangelidis, 2015; Kiratzi et al., 2016). However, the entire North Aegean Trench was activated in the west from Mount Athos to the western end of the Gulf of Saros in the east. The length of the slip zone has been estimated to be approximately 95 km longer than that predicted by empirical relationships, of the order of 50-70 km (Wells and Coppersmith, 1994). The second major earthquake occurred on June 12, 2017 with  $M_w = 6.4$  and is located approximately 15 km south of the SE coast of the Lesvos island. It was a destructive earthquake that caused one death, 15 injuries and serious damage on the island. The main shock was followed by intense aftershock activity with the strongest earthquake ( $M_w = 5.3$ ) on June 17 and its subsequent triggered seismic activity being concentrated in the eastern part of the ruptured area (Papadimitriou et al., 2018).

Another moderate event during the last decade is the 2013 January 8  $M_w = 5.8$  North Aegean earthquake sequence that took place on one of the ENE–WSW trending parallel dextral strike slip fault branches in the area, in the continuation of 1968 large ( $M = 7.5$ ) rupture (Karakostas et al., 2014). Finally, an earthquake swarm took place near the Aegean coast of NW Turkey during January–March 2017, probably related to the existence of the Tuzla geothermal field (Mesimeri et al., 2018b).





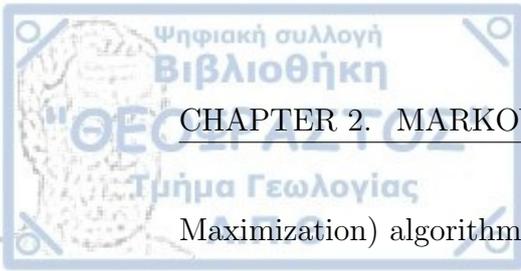
## Chapter 2

# Markovian Arrival Process modeling

### 2.1 Introduction

The MAPs are very flexible models that can approximate, given the appropriate parameterization, any complex behavior. They are extremely useful for the stochastic modelling of correlated and bursty inter-event times, as they provide a generalization of the Poisson process, PH distribution and MMPP. The development of new computational efficient methods for the estimation of the parameters led to a wider use in applied fields, especially in reliability ([Montoro-Cazorla et al., 2009](#)), queuing systems ([Lucantoni et al., 1990](#)) and internet traffic flows ([Kang et al., 2002](#)). A comprehensive review on MAPs and its applications, along with some important special cases are given in [Artalejo et al. \(2010\)](#).

First, we give the necessary mathematical background for the introduction of the MAP model along with its properties. The optimal fitting of the parameters is an open problem in the context of MAPs, so we will give a brief overview on the available methods and discuss our approach which is based on the maximization of the likelihood function through the *EM* (Expectation-



Maximization) algorithm (Dempster et al., 1977). A common difficulty in the implementation of the  $EM$  algorithm is the selection of its input values. We introduce a grid-based method for the choice of the initial parameter set which we implement in a parallel-framework for reducing the required computational time. Next, we deploy the procedure for simulating data sets with the MAP model. One basic issue when a hidden process is applied, is to estimate the most probable sequence of latent states. For this problem, we propose a local decoding algorithm with the use of the forward-backward equations (Rabiner, 1989), and we verify its stability on simulated catalogs where the state of the process at each time  $t$  is known. Finally, we present the tools for model selection and for the evaluation of the process.

## 2.2 Mathematical preliminaries

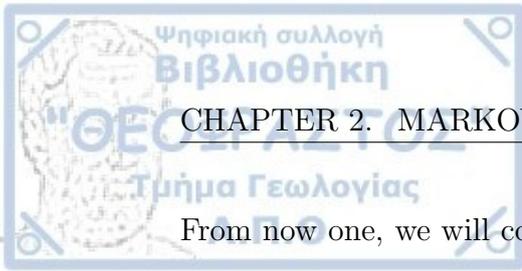
### 2.2.1 Markov Processes

The first term that we need to define is the stochastic process, i.e., a collection of random variables,  $\{X_t\}$ , with index  $t \in R^+$  in continuous time. Throughout the thesis we will study processes with discrete state space,  $\mathbf{S} = \{1, \dots, n\}$ . Now, we can give the definition of a stochastic process with the Markov property.

**Definition 2.2.1** *A stochastic process  $\{X_t : t \geq 0\}$  with state space  $\mathbf{S}$  has the Markov property when  $P(X_{t+s} = j \mid X_s = i, X_u, 0 \leq u < s) = P(X_{t+s} = j \mid X_s = i)$  holds, for each  $i, j \in \mathbf{S}$  and  $t, s > 0$  and is called Markov process.*

Essentially, the distribution of the random variable  $X_{t+s}$  is independent of the past,  $X_u, 0 \leq u < s$  and depends only on the present,  $X_s$ . The Markov process is further called homogeneous if its conditional distribution depends only on the elapsed time,  $t$ , since the current time,  $s$ ,

$$p_{ij}(t) = P(X_{t+s} = i \mid X_s = j) = P(X_t = i \mid X_0 = j).$$



From now on, we will consider only the homogeneous case. These quantities are called transition probabilities and denote the probability the process to be in state  $i$  at time  $t$  when it starts from state  $j$  at time 0. For each time  $t$  a transition probability matrix,  $\mathbf{P}(t) = \{p_{ij}(t)\}_{i,j \in \mathcal{S}}$ , is defined, whose elements are the transition probabilities among the states of the process.

The amount of time the Markov process spends in state  $i$  before moving to another state  $j$  is a random variable denoted by  $T$  that follows an Exponential distribution with parameter  $\lambda_i$ . It has therefore, the memoryless property. Next, we introduce the transition rates of a Markov process.

**Lemma 2.2.1** *For every Markov process with transition probability matrix  $\mathbf{P}(t)$  and state space  $\mathcal{S}$  the following two limits exist.*

1.  $\lim_{h \rightarrow 0^+} \frac{1-p_{ii}(h)}{h} = \lambda_i$ ,
2.  $\lim_{h \rightarrow 0^+} \frac{p_{ij}(h)}{h} = q_{ij}$  when  $i \neq j$ .

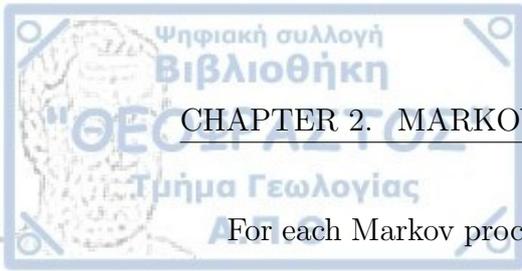
The transition rates give the infinitesimal probabilities to move among the states of the Markov process and they are represented by the matrix  $\mathbf{Q} = \{q_{ij}\}_{i,j \in \mathcal{S}}$  which is called the infinitesimal generator of the Markov process. So, the previous limits can be now written with matrix representation in the following form

$$\mathbf{Q} = \lim_{h \rightarrow 0^+} \frac{\mathbf{P}(h) - \mathbf{I}}{h},$$

where  $\mathbf{I}$  is the identity matrix. The generator of a Markov process has two main properties that are given next.

**Proposition 2.2.1** *If  $\mathbf{Q}$  is the generator of a Markov process  $X_t$  then*

1.  $\sum_{j \in \mathcal{S}} q_{ij} = 0$  for each  $i \in \mathcal{S}$ ,
2. the non-negative values of the diagonal elements of the generator matrix,  $-q_{ii} = \lambda_i$ , are the exponential parameters of the sojourn time distribution at each state  $i$ .



For each Markov process we can define a Markov Chain,  $Y_k = X_{T_k}$ ,  $k \in N$ , which is a stochastic process with the Markov property defined in discrete time. It is called embedded Markov Chain and gives at each step  $k$  the state of the Markov process at the end of the sojourn time,  $T_k$ , in the previous one. It is specified by a transition probability matrix,  $\mathbf{P} = \{p_{ij}\}_{i,j \in \mathcal{S}}$ , and an initial probability vector,  $\mathbf{a}$ , after each transition. The elements of the matrix  $\mathbf{P}$  can be easily derived by the transition rates of the generator matrix  $\mathbf{Q}$

$$p_{ij} = \frac{q_{ij}}{\sum_{j \neq i} q_{ij}} = \frac{q_{ij}}{\lambda_i}.$$

Finally, we will give the stationary probabilities of a Markov process,  $X_t$ . These are probabilities after a long time  $t$  to be at a state  $j$  that converge to a value independently from the initial state of the process.

**Definition 2.2.2** *In a Markov process  $X_t$  with state space  $\mathcal{S} = \{1, \dots, n\}$  we say*

1. *state  $j$  is reachable from state  $i$ , if  $P(X_t = j \mid X_0 = i) > 0$  for some  $t > 0$ ,*
2. *states  $i, j$  are communicating, if  $i$  is reachable from  $j$  and vice versa,*
3. *we call the process irreducible if all states communicate,*
4. *state  $i$  is called positive recurrent if the expected amount of time to return to state  $i$  given that the process started in state  $i$  is finite.*

When these conditions hold, then the stationary probabilities,  $\pi_j = \lim_{t \rightarrow \infty} p_{ij}(t)$ , exist and are derived through the linear system

$$\boldsymbol{\pi} \cdot \mathbf{Q} = \mathbf{0}, \quad \sum_{j \in \mathcal{S}} \pi_j = 1.$$

### 2.2.2. Phase-Type distributions

For the definition of this wide class of probability distributions we will follow the notation of Neuts (1978) who was the one that introduced them.

A Markov process has an absorbing state when there exists a state  $i$  that it moves to with probability equal to 1 and stays there for infinite time.

**Definition 2.2.3** *Let,  $X_t$  a Markov process with a generator matrix  $Q$ . If for a state  $i \in \mathcal{S}$ ,  $q_{ii} = 0$ , then state  $i$  is called an absorbing state.*

The transition probabilities from an absorbing state  $i$  to another state  $j$  are zero ( $p_{ij} = 0$ ).

**Definition 2.2.4** *A state  $i \in \mathcal{S}$  of a Markov process  $X_t$  is called transient if the return time probability is less than one.*

**Definition 2.2.5** *A Markov process  $X_t$  with state space  $\mathcal{S}$  is called an absorbing Markov process if all the states are transient or absorbing.*

Let us now have an absorbing Markov process,  $X_t$ , with a finite space of transient states,  $E = \{1, \dots, n\}$ , and an absorbing state,  $E_A = \{n + 1\}$ .

**Definition 2.2.6** *A Phase-Type distribution is defined as the time  $X$  that an absorbing Markov process  $X_t$  needs to move from the transient space to the absorbed one,  $n + 1$ .*

The generator matrix of the Markov process  $X_t$  has the following form

$$Q = \begin{pmatrix} D_0 & \mathbf{d}_1 \\ \mathbf{0} & 0 \end{pmatrix},$$

where  $D_0 = \{q_{ij}\}_{i,j \in E}$  is a non-singular  $(n \times n)$  transition rate matrix between the transient states and  $\mathbf{d}_1$  the  $(n \times 1)$  vector of transition rates from the transient states to the absorbing state of the process. The transient states of the Markov process  $X_t$  are called phases and their number corresponds to

the order of the Phase-Type distribution. Finally, we need to define the initial probability vector,  $\mathbf{a} = \{a_1, \dots, a_n\}$ , which gives the probability to start from a transient state  $i = 1, \dots, n$ . The probability to start from the absorbing state is neglected. So, it can be easily seen that the parameter set  $(\mathbf{a}, \mathbf{D}_0)$  specifies a Phase-Type distribution. The vector  $\mathbf{d}_1$  can be easily derived through

$$\mathbf{D}_0 \cdot \mathbf{1} + \mathbf{d}_1 = \mathbf{0},$$

as a consequence of the definition of a generator matrix, where  $\mathbf{1}$  stands for the unit column vector.

Next, we give a brief description of a PH distribution with three states and the following matrix representation

$$\mathbf{Q} = \begin{pmatrix} -1 & 0.5 & 0.5 \\ 3 & -4 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{a} = [0.6, 0.4].$$

The sojourn time at states one and two follows an Exponential distribution with parameters,  $-q_{11} = 1$  and  $-q_{22} = 4$ , respectively. At the end of the sojourn time at each phase a transition will occur either to a transient state 1, 2 with probability,  $p_{ij} = q_{ij} / -q_{ii}$ , with  $i = 1, 2$  or to the absorbing state 3 with probability,  $d_{i1} / -q_{ii}$ , where  $d_{i1}$  the  $i$ -th element of vector  $\mathbf{d}_1$ . Figure 2.1 shows the flow diagram among the phases of the Markov process.

The distribution function of a PH distributed random variable with representation  $(\mathbf{a}, \mathbf{D}_0)$  is given by

$$F(x) = 1 - \mathbf{a}e^{\mathbf{D}_0 x} \mathbf{1}, \quad x \geq 0,$$

and its density distribution is given by

$$f(x) = \mathbf{a}e^{\mathbf{D}_0 x} \mathbf{d}_1, \quad x \geq 0,$$

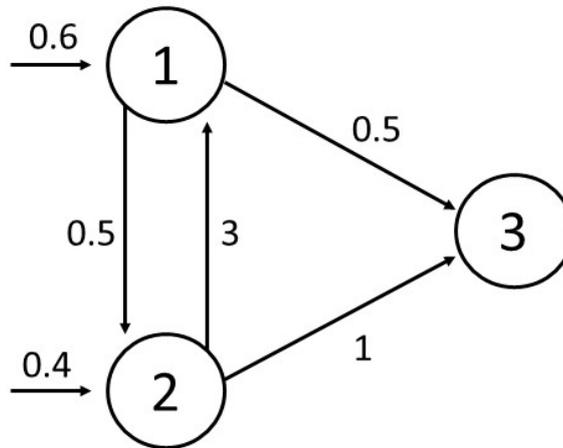


Figure 2.1: A state transition diagram of an absorbing Markov process with 2 transient and 1 absorbing state, respectively.

respectively. One of the main properties of PH distributions that makes them versatile and wide applicable in stochastic modeling is that they are dense in the class of real positive valued distributions. That is, any distribution with a strictly positive density in  $(0, \infty)$  can be approximated arbitrarily close by a PH distribution (O'cinneide, 1999).

Finally, we give some examples of special cases that can be formulated as PH distribution. The simplest one is the Exponential distribution, with a single phase and cumulative distribution function

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0.$$

In this case, the generator matrix has the following form

$$Q = \begin{pmatrix} -\lambda & \lambda \\ 0 & 0 \end{pmatrix}.$$

Another widely used special case of PH distributions is the Erlang distribution which is the sum of Exponential ones with the same parameter  $\lambda$ . Let us have  $n$  random variables  $X_i \sim Exp(\lambda)$ ,  $i = 1, \dots, n$ . Then, the random

variable  $Y = \sum_{i=1}^n X_i$  follows an Erlang distribution,  $E(n, \lambda)$ , with generator matrix

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 & 0 \\ 0 & -\lambda & \lambda & \dots & 0 & 0 \\ \vdots & \ddots & -\lambda & \dots & 0 & 0 \\ 0 & 0 & 0 & \ddots & -\lambda & \lambda \\ 0 & 0 & 0 & \ddots & 0 & 0 \end{pmatrix}$$

and cumulative function

$$F(x) = 1 - \sum_{i=1}^{n-1} \frac{(\lambda x)^i}{i!} e^{-\lambda x}, \quad x \geq 0,$$

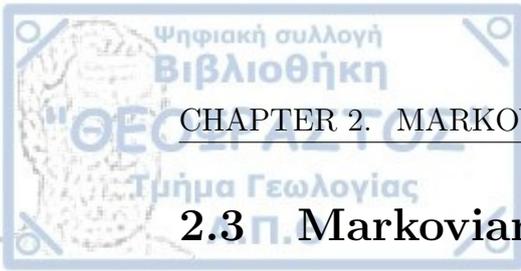
Finally, the Hyper-Exponential distribution is another widely used class of distributions that can be represented as a PH one which we will adopt in the sequel. A Hyper-Exponential (HE) distribution is a mixture of  $k$  exponentials for some  $k$  with corresponding cumulative function

$$F(x) = 1 - \mathbf{a}e^{\mathbf{D}_0 x} \mathbf{1}_k = 1 - \sum_{i=1}^k a_i e^{-\lambda_i x}, \quad x \geq 0.$$

where  $\mathbf{D}_0 = \text{diag}(-\lambda_1, \dots, -\lambda_k)$  is a diagonal matrix. The generator matrix of the PH distribution has the following form

$$Q = \begin{pmatrix} -\lambda_1 & 0 & \dots & 0 & \lambda_1 \\ 0 & -\lambda_2 & 0 & \dots & \lambda_2 \\ \vdots & \ddots & \dots & 0 & \vdots \\ 0 & 0 & \ddots & -\lambda_k & \lambda_k \\ 0 & 0 & \ddots & 0 & 0 \end{pmatrix}.$$

This class of distributions has been proven efficient in approximating long-tail probability distributions such as the Pareto and Weibull (Feldmann and Whitt, 1998) and has been widely applied in communication and network theory (Orlik and Rappaport, 1998).



## 2.3 Markovian Arrival Process

### 2.3.1 General case

Let us now consider a phenomenon where events occur at time instants,  $t_i$ , and their inter-event times,  $\tau_i = t_i - t_{i-1}$ , are correlated. Let's denote as  $N_t$  the counting process that represents the number of events up until time  $t$ . In PH distributions the absorption state can be considered as an arrival, however, the inter-event times (time until absorption  $X$ ) are independent since after each arrival the initial state is determined by the probability vector  $\mathbf{a}$ .

To overcome this issue, Neuts (1979) introduced a Markov process with two different levels of transitions. The first one concerns transitions among  $K$  states which are given by the transition rate matrix,  $\mathbf{D}_0$ . The second one, concerns transitions among the  $K$  states of the Markov process along with an arrival, which we denote  $q_{ij}(1)$  and are given through the rate matrix  $\mathbf{D}_1$ . Essentially, there are two processes running simultaneously. A counting one,  $\{N_t\}_{t \geq 0}$ , that determines the number of arrivals up to time  $t$  and a Markov process,  $J_t$ , with state space  $E = \{1, \dots, K\}$ ,  $K \in \mathbb{N}^+$  that modulates the intensity of the counting process according to its state at time  $t$ . The states of  $J_t$  are unobserved and it is called the underlying process of the MAP model. Henceforth, arrivals are translated into earthquake occurrences.

A sample path with two hidden states, initial probability vector  $\mathbf{a} = [0.8 \ 0.2]$  at time 0 and corresponding rate matrices  $\mathbf{D}_0 = \begin{pmatrix} -20 & 0 \\ 0 & -1 \end{pmatrix}$  and  $\mathbf{D}_1 = \begin{pmatrix} 17 & 3 \\ 0.25 & 0.75 \end{pmatrix}$  is presented in Figure 2.2 in order to illustrate the behavior of the model. When the underlying process,  $J_t$ , is in state 1 at time  $t = 0$  (blue line in Figure 2.2) the sojourn time of the process follows an exponential distribution with parameter  $\lambda_1 = -q_{11}(0) = 20$  and as a consequence, events occur according to a Poisson process with occurrence rate 20 (orange line in Figure 2.2). After each arrival the MAP may switch to state 2, with probabil-

ity  $q_{12}(1)/(-q_{11}(0)) = 0.15$ , or stay in the same state, state 1, with probability 0.85. These are the transition probabilities of the embedded Markov Chain. Similarly, when the process is in state 2 earthquakes are generated following a Poisson process with occurrence rate  $\lambda_2 = -q_{22}(0) = 1$ , and thus the time until the next arrival follows an exponential distribution with parameter value equal to 1. Examples with different structures of the rate matrices are given by He (2010).

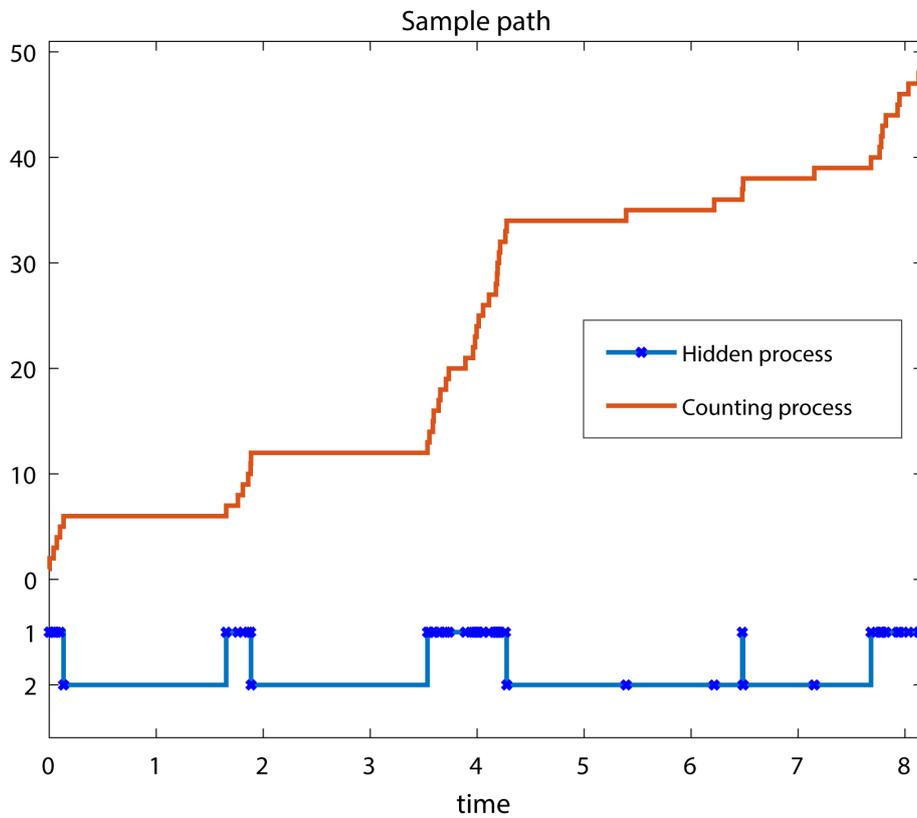


Figure 2.2: An illustration of a MAP with two states for a period (0,8.2).

Another way to visualize the two different types of transitions is through a flow diagram, given in Figure 2.3, for initial probability vector  $\mathbf{a} = [0.8 \ 0.2]$  and corresponding rate matrices  $\mathbf{D}_0 = \begin{pmatrix} -4 & 1 \\ 0.1 & -0.5 \end{pmatrix}$  and  $\mathbf{D}_1 = \begin{pmatrix} 2 & 1 \\ 0.15 & 0.25 \end{pmatrix}$ . The transitions among the same states are not allowed in Markov processes, however, introducing the second rate matrix,  $\mathbf{D}_1$ , this can be achieved along with an occurrence of an event. The transitions among the states without an

arrival, governed by rate matrix  $\mathbf{D}_0$  are shown with the thick black arrows, whereas transitions along with an arrival, governed by rate matrix,  $\mathbf{D}_1$ , are shown with the dashed arrows. It can be clearly seen the dependence of the next inter-event time to the last visited state of the underlying Markov process,  $J_t$ .

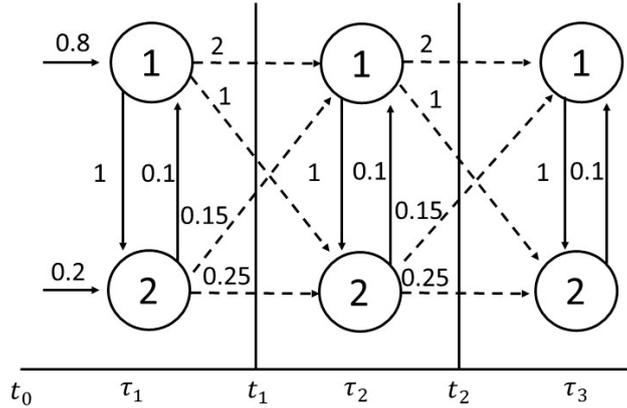


Figure 2.3: A state transition diagram of a MAP process with two states.

Hence, given the parameters  $(\mathbf{a}, \mathbf{D}_0, \mathbf{D}_1)$  the MAP can be defined in the following way.

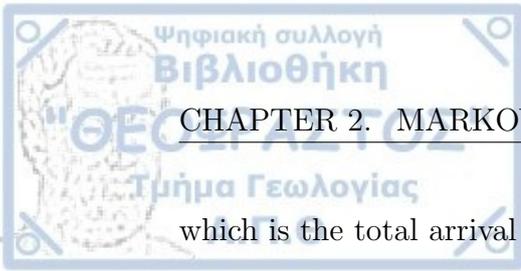
**Definition 2.3.1** A Markovian Arrival Process is a two-dimensional Markov process,  $(N_t, J_t)_{t \geq 0}$  with state space  $E = \{1, \dots, K\}$  and infinitesimal generator matrix of the underlying Markov process  $\mathbf{D}_0 + \mathbf{D}_1$ , of order  $K$ . All the elements of the two matrices are non-negative except the diagonal elements of  $\mathbf{D}_0$ , which are negative.

The generator matrix of the MAP process is an infinite block matrix with the following form

$$\mathbf{Q} = \begin{pmatrix} \mathbf{D}_0 & \mathbf{D}_1 & & \\ & \mathbf{D}_0 & \mathbf{D}_1 & \\ & & \ddots & \ddots \end{pmatrix}.$$

Every latent state of the Markov process,  $J_t$ , is linked to an occurrence rate,

$$\lambda_i = \mathbf{1}_i^T \mathbf{D}_1 \mathbf{1}_K = \sum_{j=1}^K q_{ij}(1), \quad (2.1)$$



which is the total arrival rate during the sojourn of the process at state  $i$  and is a superposition of  $K$  Poisson processes with  $q_{ij}(1)$  occurrence rates, respectively. In the sequel, we consider the stationary version of the MAP which is composed solely by parameters  $(\mathbf{D}_0, \mathbf{D}_1)$ . The initial probability vector is now denoted by  $\boldsymbol{\pi}$  and is the stationary distribution of the so called underlying or latent process,  $J_t$ . It is given through the solution of the linear system  $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$  and  $\boldsymbol{\pi}\mathbf{1}_K = 1$ . The stationary probability vector after arrival instants, denoted as  $\boldsymbol{\pi}_{arr}$ , is given by

$$\boldsymbol{\pi}_{arr} = \frac{-\boldsymbol{\pi}\mathbf{D}_0}{\boldsymbol{\pi}\mathbf{D}_1\mathbf{1}_K},$$

which is the solution of the linear system  $\boldsymbol{\pi}_{arr}\mathbf{P} = \boldsymbol{\pi}_{arr}$  and  $\boldsymbol{\pi}_{arr}\mathbf{1}_K = 1$ , with

$$\mathbf{P} = (-\mathbf{D}_0)^{-1}\mathbf{D}_1, \quad (2.2)$$

the transition probability matrix of the embedded Markov Chain of the underlying process,  $J_t$ .

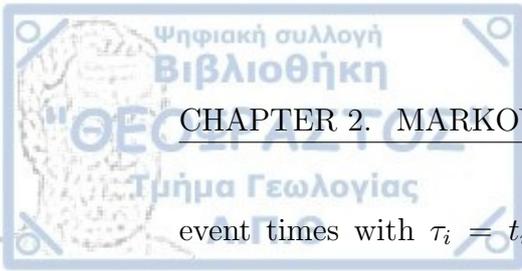
One of the most important theoretical properties of the MAP models that demonstrate their generality and versatility is due to [Asmussen and Koole \(1993\)](#), who showed that any stationary point process can be approximated arbitrarily close by a MAP model. We can see that through the parameterization of the rate matrices,  $\mathbf{D}_0$ ,  $\mathbf{D}_1$ , that can produce different types of arrival process.

One such example is the PH-renewal process ([Neuts, 1978](#)), which is a renewal process with independent inter-event times having a common PH-distribution  $(\boldsymbol{\pi}_{arr}, \mathbf{D}_0)$  after each arrival. A MAP representation is with  $\mathbf{D}_1 = \mathbf{d}_1\boldsymbol{\pi}_{arr}$ . The transitions associated with  $\mathbf{D}_0$  have no arrivals, whereas transitions associated with  $\mathbf{d}_1\boldsymbol{\pi}_{arr}$  concern arrivals.

In our case, we focus on a MAP with a diagonal matrix,  $\mathbf{D}_0 = \text{diag}(\lambda_1, \dots, \lambda_K)$ , as the example presented in Figure 2.2. The arrival rate is modulated step-wise only at the occurrence times, which is more compatible with the assumption that earthquakes are triggered by previous ones. We denote  $\lambda_1, \dots, \lambda_K$ , the non-negative diagonal elements of  $\mathbf{D}_0$ , which correspond to distinct Poisson rates and are equal to the total occurrence rate at each hidden state (Equation 2.1).

The variability of the seismicity rate in earthquake catalogs, triggered by various effects, can be captured through the transitions among different intensity rates of the counting process  $N_t$ . For instance, we know that the aftershock decay of earthquakes is following the Omori law ([Omori, 1894](#)), whereas earthquake swarms which are sequences driven by other underlying triggering mechanisms are evolving based on different temporal distributions. With MAP modeling, we can approximate both behaviors by considering two different groups of states, each one indicating a mixture of Exponential distributions for the inter-event times or equivalently a switch between Poisson rates. The transitions among states are determined through matrix  $\mathbf{D}_1$  and specify whether we are in the class that describes aftershock sequences or in the one for earthquake swarms. Generalising, the concept of embedding multiple arrival rates into the model enables the modeling of all the different temporal behaviors of earthquakes. However, it is evident that the number of states needed to approximate the temporal evolution of an earthquake catalog can be very large.

In addition, [Bountzlis et al. \(2019\)](#) showed that the tractability of the MAP due to the analytical forms of the generating functions and moments of the inter-event and arrival distributions enable to infer details on the evolution of the seismicity rate in time as well as the expected frequency of events at different time periods. Let the trace  $Tr = \{\tau_1, \dots, \tau_N\}$  be a sequence of inter-



event times with  $\tau_i = t_{i+1} - t_i$  and  $N + 1$  occurrences. The joint density function of a MAP generating  $N + 1$  consecutive events with inter-event times,  $\tau_i$ , is given by

$$f(\tau_1, \dots, \tau_N) = \boldsymbol{\pi}_{arr}^T e^{\mathbf{D}_0 \tau_1} \mathbf{D}_1 \dots e^{\mathbf{D}_0 \tau_N} \mathbf{D}_1 \mathbf{1}_K. \quad (2.3)$$

We consider the stationary version of the MAP and as initial time,  $t = 0$ , the occurrence of the first event.

The inter-event times are distributed according to a special case of the PH distribution, the Hyper-Exponential distribution, with density function

$$f(\tau) = \boldsymbol{\pi}_{arr} e^{\mathbf{D}_0 \tau} \mathbf{D}_1 \mathbf{1}_K = \sum_{i=1}^K \pi_i e^{-\lambda_i \tau} \lambda_i. \quad (2.4)$$

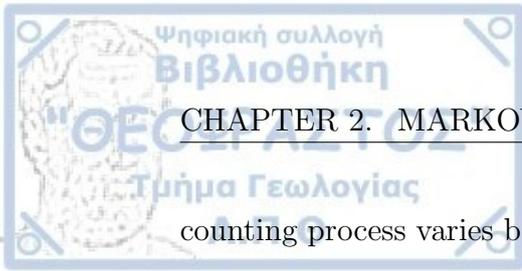
Considering a diagonal matrix for  $\mathbf{D}_0$  we derive much simpler functional forms and avoid matrix exponential computations. The embedded transition probability matrix given in Equation 2.2 can be easily derived by

$$p_{ij} = \frac{q_{ij}(1)}{\lambda_i}, \quad \forall i, j \in E. \quad (2.5)$$

More analytical measures that capture potential correlations among the events such as the joint moments of  $k$  consecutive events and the auto-correlation function between the first and  $k$ -th event are given in [Buchholz et al. \(2014\)](#).

### 2.3.3 Markov Modulated Poisson Process

A widely used special case of MAPs is the Markov Modulated Poisson Process ([Fischer and Meier-Hellstern, 1993](#)). The transition rate matrices take the following form  $\mathbf{D}_0 = \mathbf{Q} - \boldsymbol{\Lambda}$ ,  $\mathbf{D}_1 = \boldsymbol{\Lambda}$ , where  $\boldsymbol{\Lambda}$  is a diagonal matrix with positive elements. In this MAP formulation, changes in the occurrence rate can be made only between arrivals and new occurrences are generated only during the sojourn of the process in the same state. Namely, on MMPP the rate of the

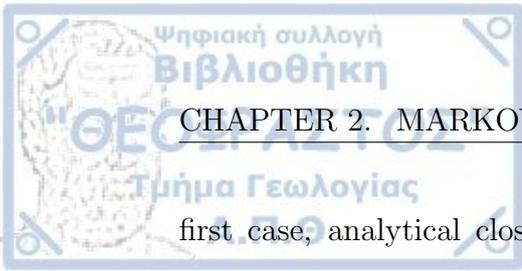


counting process varies between the occurrence of two events according to the underlying Markov process,  $J_t$ , but transitions are not allowed at an occurrence instant, contrary to our case, where changes in the occurrence rate are allowed only when an earthquake occurs. Thus, the time between the  $(k-1)$ -th and  $k$ -th events depends on the state of  $J_t$  at the  $(k-1)$ -th occurrence, whereas on a MMPP, the underlying process,  $J_t$ , can move several times to different hidden states, following the  $k$ -th arrival with a transition to the same state. From a seismological point of view, we assume that changes in the seismicity rate can be triggered by a previous earthquake, while under the MMPP formulation seismicity rate can change multiple times before the earthquake occurrence. We believe that these conditions have an impact on the adaptability of the model to the data, as the structure of the infinitesimal generator changes.

## 2.4 Parameter estimation

Fitting of MAPs is the most challenging part for real data applications mainly due to the large number of free parameters and the lack of canonical representations. The former increases the complexity since we need large data sets to capture the correlations among the events. In addition, the parameters of the model are not uniquely determined, meaning that the same likelihood function can be derived by at least two different parameter sets, known as identifiability problem. This explains the multimodality of the likelihood function which consists one of the most significant problems for the selection of the parameters. However, recent efforts have successfully found unique representations for different classes of MAP models (Bodrog et al., 2008; Telek and Horváth, 2007).

In general, there are two main approaches for the estimation of the MAPs parameters. Moment matching methods, where the theoretical moments should be equal to the observed ones and likelihood-based methods where the parameter set corresponds to the maximum value of the likelihood function. In the



first case, analytical closed forms exist for the two-state MAP with hyper-exponential inter-event times distribution (Heindl, 2004). However, the spectrum of states that the method can be applied is limited. As their number increases the higher order or joint moment equations become non-linear and optimization methods do not work. There are although some compositional efforts that combine the fitting capability of lower order MAPs with the flexibility of ones with more states. One such effort is by Casale et al. (2010) who proposed the Kronecker Product Composition method, where MAPs are constructed by processes of a smaller number of states through Kronecker products. Another popular framework consists of a two-phase fitting procedure. At the first step, a PH distribution with parameters  $(\boldsymbol{\pi}, \mathbf{D}_0)$  is fitted by any available method to the data and at the second phase the elements of the transition matrix  $\mathbf{D}_1$  are estimated by various methods, such as first (Buchholz and Kriege, 2009) and higher lag joint moments (Bause and Horváth, 2010) or matching the lag- $k$  auto-correlation function (Horváth et al., 2005).

There are several Maximum Likelihood Estimation (MLE) fitting methods that are adopted for the MAP transition rates. The general form of the likelihood function is given through Equation (2.3),

$$L(\boldsymbol{\theta}|Tr) = \boldsymbol{\pi}_{arr}^T e^{\mathbf{D}_0 \tau_1} \mathbf{D}_1 \dots e^{\mathbf{D}_0 \tau_N} \mathbf{D}_1 \mathbf{1}_K, \quad (2.6)$$

with parameter set  $\boldsymbol{\theta} = \{q_{ij}(0), q_{ij}(1)\}$ , trace  $Tr$ ,  $N + 1$  events and  $K$  latent states. It considers the whole trace in order to capture the correlations among the events. The optimization of the function,  $\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\boldsymbol{\theta}} L(\boldsymbol{\theta}|Tr)$ , is based on the implementation of an iterative algorithm, the *EM* algorithm (Dempster et al., 1977), which is an appropriate technique especially in problems with unobserved data and it consists of the E-step (Expectation) and M-step (Maximization).

If  $\boldsymbol{\theta}_0$  is a given initial parameter set, then in the E-step random variables are created related to the unobserved data  $T'$  and an expected log-likelihood

function ( $LL$ ),  $E[\log L(\boldsymbol{\theta} \mid Tr, T') \mid Tr; \boldsymbol{\theta}_0]$ , is computed conditional on the complete data vector  $(Tr, T')$ . In our case, the variables are related to the transitions among the latent states of the process,  $J_t$ . In the M-step, the new parameter set  $\boldsymbol{\theta}_1$  is estimated through maximizing the expected  $LL$  function,  $\boldsymbol{\theta}_1 = \operatorname{argmax}_{\boldsymbol{\theta}} E[\log L(\boldsymbol{\theta} \mid Tr, T') \mid Tr; \boldsymbol{\theta}_0]$ . At each iteration of the algorithm the likelihood function is expressed through the forward and backward vectors. The  $i$ -th element of the forward row vector  $\mathbf{f}[k] = \{f_i(k), i = 1, \dots, K\}$  denotes the likelihood to be in state  $i$  conditional on the history of the process up to the occurrence time  $t_{k+1}$  and can be obtained recursively by  $\mathbf{f}[0] = \boldsymbol{\pi}$ ,  $\mathbf{f}[k] = \mathbf{f}[k-1]e^{D_0\tau_k} \mathbf{D}_1$ . Similarly, we can define the backward vectors  $\mathbf{b}[N+1] = \mathbf{1}_K$ ,  $\mathbf{b}[k] = e^{D_0\tau_k} \mathbf{D}_1 \mathbf{b}[k+1]$ , with  $k = 1, \dots, N$ . The algorithm ends when a convergence criterion is satisfied or a certain number of iterations is applied.

[Asmussen et al. \(1996\)](#) and [Rydén \(1996\)](#) incorporated the algorithm into the framework of PH distributions and MMPPs, respectively. Their innovative work was followed by [Buchholz \(2003\)](#) who enhanced the computational speed for general MAPs through the uniformization method for the matrix-exponential computations, whereas [Okamura and Dohi \(2009\)](#) and [Breuer and Kume \(2010\)](#) worked on applying EM algorithm on grouped data of disjoint intervals, since many times real data is provided in groups, especially on network traffic data. A comprehensive survey on the most recent state-of-art results on MAP fitting methods can be found in [Buchholz et al. \(2014\)](#) and [Okamura and Dohi \(2016\)](#). However, the implementation of the EM algorithm demands a large number of iterations, each one with high computational effort due to the evaluation of matrix-exponentials for the likelihood function. [Okamura and Dohi \(2009\)](#) suggested that the fitting with a sub-class of MAPs will reduce the computational effort. They proposed a structure with components following PH and as a special case Erlang distribution. They found that the time until convergence is being reduced significantly while the likelihood values remained high, whereas [Horváth and Okamura \(2013\)](#) have extended the method to multiclass MAPs, keeping the same special structure.

In our study, we need to fit the parameters of a simpler MAP structure described in 2.3.2. We assume that the marginal distribution of the inter-event times is given by a mixture of exponential ones, namely a Hyper-Exponential distribution. Given that the MAP is in state  $i$  just after an earthquake occurrence, the density function of the inter-event times distribution is given by

$$f_i(\tau) = \lambda_i e^{-\lambda_i \tau},$$

following Equation (2.4). Due to the special structure of the matrix,  $\mathbf{D}_0$ , the forward and backward vectors and as a sequence the likelihood function can be expressed in a simpler way

$$\mathbf{f}[k]_j = \sum_{i=1}^K f[k-1]_i e^{-\lambda_i \tau_k} q_{ij}(1), \quad \mathbf{b}[k]_j = \sum_{i=1}^K e^{-\lambda_j \tau_k} q_{ji}(1) \mathbf{b}[k+1]_i, \quad (2.7)$$

and

$$L(\boldsymbol{\theta} | Tr) = \mathbf{f}[k] \mathbf{b}[k+1], \quad k = 0, \dots, N,$$

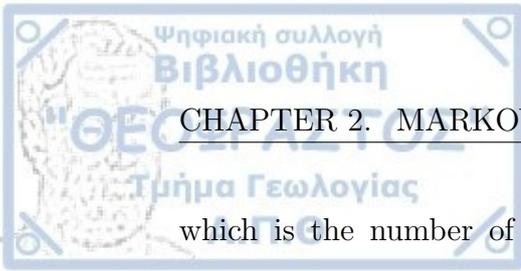
respectively, with parameter set  $\boldsymbol{\theta} = \{\lambda_i, q_{ij}(1)\}$ . In this way, we avoid the computation of matrix-exponentials (see Equation (2.6)) reducing the computational cost for the evaluation of the likelihood function.

Considering the unobserved data,  $T' = \{s_1, \dots, s_N\}$ , to be known, i.e., the hidden states at each occurrence time, we can easily obtain the maximum likelihood estimates of the parameter set  $\boldsymbol{\theta}$ . The inter-event times are conditionally independent, therefore,

$$\log L(\lambda_1, \dots, \lambda_K | Tr, T') = \sum_{k=1}^N \log(f_{s_k}(\tau_k)),$$

and considering its derivative for each  $\lambda_i$  we can obtain its maximum value

$$\hat{\lambda}_i = \frac{\sum_{k=1}^N I_{(s_k=i)}}{\sum_{k=1}^N \tau_k I_{(s_k=i)}}, \quad (2.8)$$



which is the number of intervals that the process is in state  $i$  divided by the total sojourn time in state  $i$ . The maximum likelihood estimates of the transition probabilities,  $p_{ij}$ , of the MAP's embedded Markov Chain at arrival instants, are obtained by [Anderson and Goodman \(1957\)](#) and are given by

$$\hat{p}_{ij} = \frac{\sum_{k=1}^{N-1} I_{(s_k=i, s_{k+1}=j)}}{\sum_{k=1}^{N-1} I_{(s_k=i)}}. \quad (2.9)$$

Then, the elements of the transition rate matrix,  $\mathbf{D}_1 = \{q_{ij}(1)\}_{i,j \in E}$ , are easily obtained from Equation (2.5).

In the E-step of the EM algorithm the expected values of the unobserved variables that correspond to the latent states that generate the  $N + 1$  earthquake occurrences, are derived and are expressed through the forward and backward vectors. In particular,

$$\begin{aligned} q_{ij}[k] = P(s_k = i, s_{k+1} = j | \boldsymbol{\theta}, Tr) &= \frac{P(s_k = i, s_{k+1} = j, Tr | \boldsymbol{\theta})}{P(Tr | \boldsymbol{\theta})} \\ &= \frac{\mathbf{f}[k]_i \cdot f_i(\tau_k) \cdot p_{ij} \cdot \mathbf{b}[k+1]_j}{P(Tr | \boldsymbol{\theta})}, \end{aligned}$$

and

$$q_i[k] = \frac{P(s_k = i, Tr | \boldsymbol{\theta})}{P(Tr | \boldsymbol{\theta})} = \frac{\mathbf{f}[k-1]_i \cdot \mathbf{b}[k]_i}{P(Tr | \boldsymbol{\theta})}.$$

Then, in the M-step the expected values are maximized (Equations (2.8) and (2.9)) and give the estimated parameters of set  $\boldsymbol{\theta}$ ,

$$\hat{\lambda}_i = \frac{\sum_{k=1}^N q_i[k]}{\sum_{k=1}^N \tau_k q_i[k]} = \frac{\sum_{k=1}^N \mathbf{f}[k-1]_i \mathbf{b}[k]_i}{\sum_{k=1}^N \tau_k \mathbf{f}[k-1]_i \mathbf{b}[k]_i}, \quad (2.10)$$

and

$$\hat{p}_{ij} = \frac{\sum_{k=1}^{N-1} q_{ij}[k]}{\sum_{k=1}^{N-1} q_i[k]} = \frac{\sum_{k=1}^N \mathbf{f}[k-1]_i e^{-\lambda_i \tau_k} q_{ij}(1) \mathbf{b}[k+1]_j}{\sum_{k=1}^N \mathbf{f}[k-1]_i \mathbf{b}[k]_i}. \quad (2.11)$$

Finally, the transition rates,  $q_{ij}(1)$ , are easily obtained by,  $q_{ij}(1) = \hat{\lambda}_i \cdot \hat{p}_{ij}$ .

### 2.4.1.1. Determination of the initial values

For the implementation of the EM algorithm, BuTools program package (Bodrog et al., 2014) is used in the MATLAB environment. The algorithm requires as input the trace and an initial choice for the parameter vector. However, as the number of parameters is increasing, the  $LL$  shows more local peaks, making it more difficult to converge to the global maximum. The dependency on the choice of the initial parameter values is also increasing as the length of the sample is decreasing.

Therefore, we implemented a grid-based procedure for the initial choice that it consists of the following steps:

- Construct grid for the occurrence rates,  $\lambda_i$ ,  $i = 1, \dots, K$ , ranging from  $N_{max}^{obs}$  to  $N_{min}^{obs}$  with a constant step,  $s$ , leading to  $N_{st} = \frac{N_{max}^{obs} - N_{min}^{obs}}{s}$  values. The first is the maximum observed number of events in the time unit and the second the minimum one, respectively. Divide it into  $K$  intervals equal to the number of hidden states and consider all the possible combinations. This equals to  $N_{all} = (N_{st}/K)^K$  sets of values.
- Generate the transition probabilities,  $p_{ij}$ , uniformly under ergodicity conditions. For each set  $\{\lambda_1^j, \dots, \lambda_K^j\}$ , with  $j = 1, \dots, N_{all}$ , correspond  $N$  random transition probability matrices,  $\mathbf{P}$ .
- The total  $LL$  computations for the choice of the initial vector:  $N \times N_{all}$ .

The choice of  $N$  is independent of the occurrence rates grid length. Therefore, we can reduce the computation time by implementing a parallel procedure. Let  $M$  be the number of parallel workers. Then, we can divide the initial grid into  $M$  groups leading to  $N_{all}/M$  of  $\{\lambda_1, \dots, \lambda_K\}$  sets in each one. The rest of the procedure is the same, leading to  $N \times (N_{all}/M)$  final computations. Finally,  $N_{best}$  memory storage is needed for the best parameter sets in terms of their log-likelihood values which are given then as input to the EM algorithm.

Concerning the complexity of the EM algorithm, essentially we need to compute and store the forward and backward vector,  $\mathbf{f}[k]$  and  $\mathbf{b}[k]$ , respectively, for  $k = 1, \dots, N$ . This requires  $2N$  vector-matrix multiplications of size  $K$ , which is the number of latent states, and a  $2N$  memory storage, respectively. Recently, 3 refined algorithms for the parallel implementation of the EM algorithm have been developed by [Bražėnas et al. \(2018\)](#), that showed very good performance on numerical experiments.

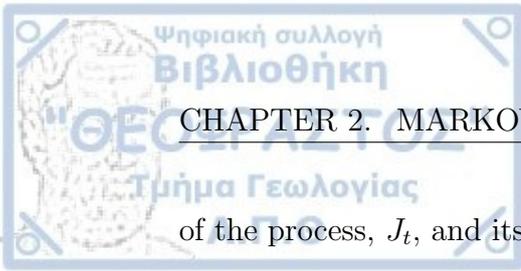
## 2.5 Simulation procedure

We want to create a simulated sequence of events that is generated by the counting process,  $N_t$ , based on the rate matrices,  $\mathbf{D}_0 = -diag(\lambda_1, \dots, \lambda_K)$ ,  $\mathbf{D}_1 = \{q_{ij}(1)\}_{i,j \in E}$  and the stationary probability vectors  $\boldsymbol{\pi}_{arr}$  and  $\boldsymbol{\pi}$  depending on whether an arrival occurred at the initial point  $t_0 = 0$  or not.

To simulate a MAP, first we need to determine the hidden state of the process at time  $t_0 = 0$ . Select an initial probability vector, in our case, a stationary vector and generate the initial state  $J(t_0)$  with the use of a multinomial trial from the discrete probability distribution  $\boldsymbol{\pi}_{arr}$  or  $\boldsymbol{\pi}$ . Then, the sojourn of the process at each state and the type of arrival need to be determined. In our case, we have only a single type of arrival since  $\mathbf{D}_0$  is a diagonal matrix, namely, transitions among the hidden states occur only along with an earthquake occurrence. So, sojourn times,  $\tau$ , are generated from Exponential distributions with parameters  $\{-\mathbf{D}_0\}_{ii} = \lambda_i$  under the relationship

$$\tau = -\frac{\ln U}{\lambda_i}, \quad (2.12)$$

that is derived by a typical thinning procedure. At the end of the sojourn time at each state we need to determine the next hidden state of the process,  $J_t$ . This is resolved by the conditional probabilities of the MAP,  $\mathbf{P}(J(T_0 + \tau)/J(T_0)) = \{p_{ij}\}_{i,j \in E}$  where  $\mathbf{P}$  is the embedded transition probability matrix



of the process,  $J_t$ , and its elements are given by Equation (2.5). In particular, the next state  $j$  is generated by a multinomial trial from the discrete probability distribution

$$\left\{ \frac{q_{i1}(1)}{\lambda_i}, \dots, \frac{q_{iK}(1)}{\lambda_i} \right\},$$

under the condition that the previous state is  $i$ . The procedure is repeated until a certain number of events is derived or a time  $t_{end}$  is exceeded.

## 2.6 Inference of the latent states

One of the three basic issues related to the HMMs and as a consequence to the MAPs is the decoding problem. In other words, to find the optimal path of latent states that have generated the sequence of observations,  $t_1, \dots, t_{N+1}$ , given the parameter set,  $\boldsymbol{\theta} = \{\lambda_i, q_{ij}\}$ . From a set of probabilities we choose the one with the maximum value, i.e.,

$$J_{t_k} = \operatorname{argmax}_{1 \leq i \leq K} P(J_{t_k} = i | Tr, \boldsymbol{\theta}), \quad (2.13)$$

which is known as local decoding problem. To solve this kind of problem we need to follow the corresponding formulation for the HMMs from [MacDonald and Zucchini \(1997\)](#). In particular, the probabilities given in Equation (2.13), which we will call state probabilities, can be written

$$P(J_{t_k} = i | Tr, \boldsymbol{\theta}) = p_i(t_k) = \frac{P(\tau_1, \dots, \tau_N, J_{t_k} = i)}{P(\tau_1, \dots, \tau_N)},$$

through a simple implementation of Bayes theorem. Then, we use the forward and backward vectors defined by Equation (2.7) to evaluate these probabilities. Both of them are crucial as the first,  $\mathbf{f}[k]_i$ , gives the probability the state at time  $t_k$  to be  $i$  and  $k$  earthquakes to have occurred with inter-event times,  $\tau_1, \dots, \tau_k$ . Conversely, the backward vector,  $\mathbf{b}[k]_i$ , is the probability of observing  $N + 1 - k$  earthquakes with inter-event times,  $\tau_{k+1}, \dots, \tau_N$ , and the

latent state,  $i$ , at time  $t_k$ . The state probabilities therefore, take the following form

$$p_i(t_k) = \frac{\mathbf{f}[k-1]_i \cdot \mathbf{b}[k]_i}{L(\boldsymbol{\theta} | Tr)} = \frac{\mathbf{f}[k-1]_i \cdot \mathbf{b}[k]_i}{\sum_j^K \mathbf{f}[k]_j \cdot \mathbf{b}[k+1]_j}. \quad (2.14)$$

Due to the large number of matrix multiplications for the computation of the forward and backward vectors it is quite often to have overflow and underflow problems, so the vectors  $\mathbf{f}[k]$  and  $\mathbf{b}[k]$  are scaled accordingly (Bražėnas et al., 2018).

Finally, since the MAP model can be defined as a stochastic point process we can also evaluate the intensity function,  $\lambda(t)$ , of the counting process,  $N_t$ , expressed by

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{P_{dt}(t/H_t)}{dt}.$$

Its estimator, can be easily derived through the forward and backward vectors as follows

$$\hat{\lambda}(t) = \sum_{i=1}^K \hat{\lambda}_i \cdot \hat{p}_i(t_k) = \sum_{i=1}^K \frac{\sum_{k=1}^N \mathbf{f}[k-1]_i \mathbf{b}[k]_i}{\sum_{k=1}^N \tau_k \mathbf{f}[k-1]_i \mathbf{b}[k]_i} \cdot \frac{\mathbf{f}[k-1]_i \cdot \mathbf{b}[k]_i}{\sum_j^K \mathbf{f}[k]_j \cdot \mathbf{b}[k+1]_j}, \quad (2.15)$$

for  $t_k \leq t < t_{k+1}$ .

### 2.6.1 Stability of the local decoding algorithm

The stability of the state probability estimates given in Equation 2.14 is tested on simulated data sets from a MAP model with four states where the hidden path is known a-priori. In more detail, four samples with  $N_k = 50, 100, 200$  and 300 events are simulated from a MAP with four states along with the path of the hidden states (thick grey line in Figure 2.4). Then, a new MAP is fitted based on the trace, the inter-event times, of each sample for 100 times, and the sequence of the latent states that have generated the events is re-evaluated for each fitted MAP though Equations 2.13 and 2.14. The average estimated state at each occurrence time is computed and in Figure 2.4 the simulated hidden path is shown with the grey color and the average estimated hidden

path over the 100 estimated hidden sequences with the dash-dotted black line.

We note, that the average value does not have a physical meaning, since the states are integer values, however, it can be seen as an index of how close to the simulated path our estimation is after 100 implementations. We can see that the average values approach the true sequence with high accuracy especially as the number of events increases.

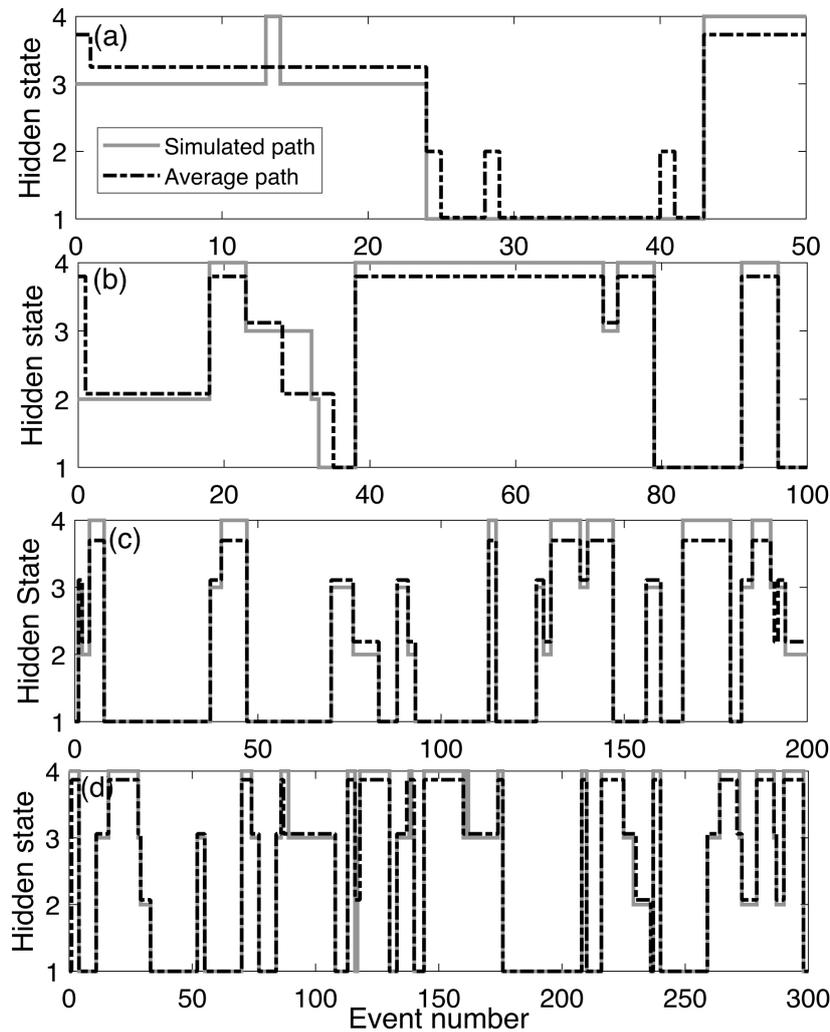


Figure 2.4: The MAP is fitted to simulated samples of a four-state MAP with  $N_k = 50, 100, 200$  and  $300$  events for 100 times. The line in gray color corresponds to the hidden path of each simulated sample. Next, the sequence of the underlying states that have generated the events is evaluated (100 times) and the average estimated state for each occurrence time is computed (dash-dotted black line).

When we want to approximate empirical data, in our case earthquake data sets, with the use of the MAP model, first we need to select the appropriate number of states,  $K$ , of the Markov process,  $J_t$ . Due to its property, that can approximate any stationary stochastic point process (Asmussen and Koole, 1993), increasing its state space the model's theoretical moments are converging to the empirical ones. However, this can lead to overfitted models with large estimated variances exhibiting poor forecasting results. Conversely, choosing few states might lead to underfitted models that miss important features of the data.

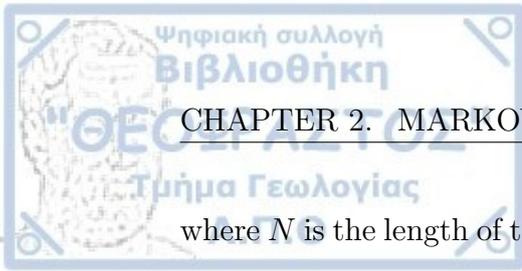
This means that the selection of the appropriate number of states is crucial for a robust stochastic modeling and for the statistical inference of the empirical data. Here, for the selection of the optimal model we will use two criteria that are based on information theory with desirable statistical properties (Burnham and Anderson, 2002). The first one is the Akaike Information Criterion (AIC) (Akaike, 1974) which is based on the maximum value of the log-likelihood function and has the following form

$$AIC = -2 \cdot LL + 2 \cdot K, \quad (2.16)$$

where  $LL$  is the maximum log-likelihood value and  $K$  the number of the model's free parameters. Essentially, we achieve a balance between errors due to underfitting and overfitting. The penalized criterion is computed for all the fitted models and the one with the lowest value is preferred as the best fit to the data without, however, quantifying the level of the fitting.

The second criterion that we will use throughout the study is the Bayes Information Criterion (BIC) (Schwarz, 1978) which additionally considers the length of the sample. Its form is

$$BIC = -2 \cdot LL + K \cdot \ln N, \quad (2.17)$$



where  $N$  is the length of the sample. Again, the model with the minimum value among the candidates is selected as optimal. The difference of the two criteria lies on the coefficient multiplying the number of parameters, in other words, on the penalized effect. [Vrieze \(2012\)](#) showed that the BIC is consistent in selecting the true model when this model is considered as a candidate (among other assumptions), however, he concludes that even if the true model is under consideration, the BIC is not a clear choice. An extended review concerning selection criteria related to maximum likelihood models is given in [Kadane and Lazar \(2004\)](#).

## 2.8 Evaluation

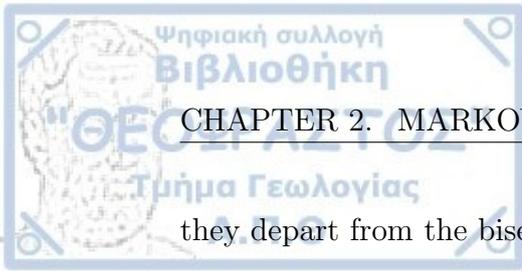
Information criteria are useful when we want to perform a comparison among competing models. However, they do not tell us how well the fitted models approximate the target data set. A useful technique for evaluating the point process goodness-of-fit is the residual analysis ([Ogata, 1988](#)), which is a time transformation test.

Let  $\{t_1, \dots, t_N\}$  be a set of occurrence times generated by a stochastic point process  $\mathbf{N}^*$  with intensity function  $\lambda(t)$ . Let also consider the integral of the intensity function

$$\tau^* = \Lambda(t) = \int_0^t \lambda(t) dt, \quad (2.18)$$

where  $\Lambda(t)$  is a monotonic increasing function of  $t$ . [Papangelou \(1972\)](#) showed that with probability 1, the sequence of transformed times,  $\tau_i^* = \Lambda(t_i)$ ,  $i = 1, \dots, N$ , is a realization of a Poisson process with a unit rate. Hence, if the fitted point model approximates well the empirical data, we expect the transformed times,  $\tau_i^*$ , derived by the estimated intensity function,  $\hat{\lambda}(t)$ , will follow a Poisson distribution with arrival rate equal to one.

An advantage of the test is that it can be visualized. In particular, we can compute the cumulative number of the transformed times,  $\tau_i^*$ , and see whether



they depart from the bisector,  $y = x$ , which is the typical behaviour of events randomly distributed in time with rate equal to one. Daley and Vere-Jones (2003) proposed an algorithm for testing the goodness-of-fit of a point process based on the residual analysis that consists of the following 4 steps.

- Compute the sequence of transformed times,  $\tau_i^*$ , through Equation (2.18).
- Plot the cumulative step-function through the pairs  $(x_i, y_i) = (\tau_i^*/T, i/N)$  with  $i = 1, \dots, N$ , and  $T$  the total duration.
- Plot the confidence bounds  $y_{int} = x \pm \frac{z_{\alpha/2}}{\sqrt{T}}$ , where  $z$  is distributed according to the standard normal distribution,  $N(0, 1)$ .
- Observe whether the transformed data lies within the  $100(1 - \alpha)\%$  confidence interval drawn in the previous step.

There are also a number of hypothesis tests that can be implemented to test whether the transformed times follow a Poisson distribution with known parameter,  $\lambda_{pois} = 1$ . Considering the inter-event transformed times,  $E_i = \tau_i^* - \tau_{i-1}^*$ ,  $i = 1, \dots, N$ , we can apply two types of test to verify the assumption of a stationary Poisson process.

The first one concerns the assumption of independence among the events. The Runs test (Bradley, 1968) is used to answer to the question on whether a sample data is generated from a random process or not. We put the values on a line and mark the ones above the median as positive whereas the ones below as negative. A run is defined as a series of consecutive positive (or negative) values and the two alternatives are defined as:

$H_0$ : the sequence of events is distributed randomly

$H_1$ : the sequence of events is not distributed randomly.

The test statistic is

$$Z = \frac{R - \bar{R}}{s_R},$$

where  $R$  is the observed number of runs,  $\bar{R}$ , is the expected number of runs, and  $s_R$  is the standard deviation of the number of runs. The values of  $\bar{R}$  and

$s_R$  are computed through the following formulas

$$\bar{R} = \frac{2n_1n_2}{n_1 + n_2} + 1, \quad s_R = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)},$$

where  $n_1$  and  $n_2$  the number of positive and negative values in the sample, respectively. An alternative technique for investigating possible correlations among the events is through the auto-correlation function (Box et al., 2015), which is defined as

$$\rho_k = \frac{\sum_{i=1}^{N-k} (E_i - \bar{E})(E_{i+k} - \bar{E})}{\sum_{i=1}^N (E_i - \bar{E})^2},$$

for lag  $k$ . If its values lies within the 95% confidence bounds there is no evidence that the inter-event times,  $E_i$ , are not independent.

The next type of test is the Kolmogorov–Smirnov (KS) test (Chakravarti et al., 1967) which is applied on the sample  $E_i$  to verify that it follows an Exponential distribution,  $Exp(\lambda = 1)$ . It computes the differences between the empirical and exponential cumulative distribution over all the sample and the test statistic is defined as the maximum distance among the two curves. An important limitation of the test is the requirement the parameters of the distribution to be fully specified instead of being estimated from the data.

When the parameters of the null distribution are not known and need to be estimated from the sample, the Lilliefors goodness of fit test (Lilliefors, 1969) is the appropriate one to use. It is similar to the KS test, however, the parameters of the null hypothesis distribution,  $F_0$ , are estimated from the sample and the  $p$  values are computed through Monte Carlo simulations since the distribution of the test-statistic is not fully defined.



## Chapter 3

# Markovian Arrival Process for earthquake clustering

### 3.1 Introduction

Earthquake clustering is an essential feature of seismicity that provides crucial information on the earthquake dynamics, expressed as the concentration of earthquakes in space and time. There are various forms of triggering mechanisms responsible for the different types of clusters including static and dynamic stress changes induced from previous events (main shock-aftershocks) (Felzer and Brodsky, 2006; Stein, 1999), fluid migration and aseismic slip (earthquake swarms) (Yamashita, 1999).

The improvement of seismic monitoring worldwide and the development of new powerful algorithms for earthquake detectability (Ross et al., 2019) increased the amount of seismic data that is available for data processing, motivating the development of more refined clustering algorithms. Earthquake clusters can be used for the determination of faulting geometry (Kamer et al., 2020; Petersen et al., 2021) as well as to extract useful information on the aftershock productivity of the study area and the behavior of the foreshock activity (Lippiello et al., 2012; Shebalin et al., 2020). However, many studies

## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

---

focus on the data after the removal of clustered events, known as background seismicity and the procedure as declustering. Background seismicity is used to Probabilistic Seismic Hazard Analysis (PSHA) for the production of hazard maps (Gerstenberger et al., 2020; Petersen et al., 2018) and can be also used to infer physical properties of main shocks such as the localization of seismicity before large earthquakes by Ben-Zion and Zaliapin (2020).

In the first part of this chapter, we establish a two-step clustering procedure that comprises a temporal stochastic point process, the Markovian Arrival Process, for an initial separation of the background seismicity from potential seismic excitations, using the changes in the seismicity rate, and the density-based clustering algorithm, DBSCAN, for the detection of elevated density areas in space. We assume that the physical mechanisms governing the earthquake clustering are unknown and the prevailing parameter to separate the background seismicity from seismic excitations are the temporal variations of the seismicity rate. For this reason, the MAP,  $(N_t, J_t)_{t \in R^+}$  is used, whose intensity function,  $\lambda(t)$ , is modulated by the latent Markov process,  $J_t$ . The earthquakes temporal distribution is essentially approximated by a non-homogeneous Poisson process with a piece-wise constant intensity rate determined by the underlying Markov process,  $J_t$ . The sequence of hidden states is evaluated through a local decoding algorithm, and the level of the occurrence rate at each time  $t$  is revealed. In this way, the MAP can be used as a tool for change point detection, namely to detect changes in the seismicity rate. We show here that the model is suitable for capturing prevailing patterns of the seismicity dynamics, by corresponding the estimated rates to known foreshock-main shock-aftershock sequences and swarms, as well as to periods of relative seismic quiescence. Recent works by Lu (2019) and Benali et al. (2020) are based on non-stationary Poisson models whose rate is modulated by a hidden Markov process to determine a set of change-points for seismicity rate. Subsequently, the DBSCAN (Ester et al., 1996) mentioned before is used for grouping the events into spatiotemporal clusters.



## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

---

In the second part of this chapter we present a comprehensive analysis of the clustering properties in three major seismic zones of Greece. Several studies suggest that the clustering properties of seismicity (spatiotemporal distribution, productivity rates) might be controlled by the tectonic regime. [Llenos and Michael \(2017\)](#) showed that the adoption of region-specific aftershock parameters can improve forecast estimates, as the information from the tectonic region is particularly useful, and suggest the determination of clustering features in smaller regions where high-quality earthquake data are available. The global analysis of earthquake clusters from [Zaliapin and Ben-Zion \(2016\)](#) suggests that seismicity clusters in a region strongly depend on the heat flow and believe that considering region specific deviation can improve local seismic hazard assessment. More recently, [Hardebeck et al. \(2019\)](#) updated the generic parameters of sequences in California incorporating the regionalization of the former work for their determination. In this way, there was an improvement of the aftershock forecasts accuracy. The temporal ETAS model assumes that background events occur according to a stationary Poisson process with rate,  $\mu$ , which can trigger other earthquakes and in turn can trigger more earthquakes and so on. In our study, we utilize the estimated parameters of the ETAS model to investigate regional variabilities in the productivity of the seismic sequences and gain insights into the involved triggering mechanisms ([Crespo Martín and Martín-González, 2021](#); [Hainzl and Ogata, 2005](#); [Marsan et al., 2013](#)).

We will thoroughly discuss the features of multiple embedded occurrence rates that can separate potential seismic clusters from background seismicity, generalized into the idea of MAP to be used as a change point tool for seismicity rates. Then, we establish a two-step clustering algorithm that combines the MAP model with a density based clustering algorithm, DBSCAN, and we evaluate its performance on a simulated earthquake catalog where the structure of the clusters is known a-priori. Finally, we apply the method to the three major seismic zones of Greece as already mentioned and investigate their



clustering properties. In particular, we focus on the statistical analysis of the detected clusters based on the ETAS model producing generic and sequence specific parameters for each area.

## 3.2 MAP-DBSCAN method

### 3.2.1 MAP as a tool for change point detection

The temporal distribution of seismicity can be approximated by a stochastic point model, the Markovian Arrival Process. Following the notation given in Section 2.3, the MAP is a two-dimensional Markov process  $(N_t, J_t)_{t \in R^+}$ , where  $N_t$  counts the number of earthquakes that occur up to time  $t$  with a rate  $\lambda_{J_t}$ , which is associated with the hidden states  $i = 1, \dots, K$ , of the Markov process,  $J_t$ . For the representation of the MAP model, we need the  $K \times K$  rate matrices  $\mathbf{D}_0$  and  $\mathbf{D}_1$ , where  $\mathbf{D}_0$  is a diagonal matrix whose non-negative elements we denote as,  $\lambda_1, \dots, \lambda_K$ , and correspond to  $K$  Poisson rates, each one assigned to a hidden state of process  $J_t$ , and  $\mathbf{D}_1$  consists of the transition rates among the states along with the occurrence of an earthquake, which we denote as  $q_{ij}(1)$ .

A brief description of the process is given next. When the Markov process  $J_t$  is in state  $i$ , earthquakes occur according to a Poisson process with seismicity rate  $\lambda_i$  and, therefore, the sojourn time in this state follows an Exponential distribution with expected value  $1/\lambda_i$ . When an earthquake occurs, the MAP can move with transition rate  $q_{ij}(1) = p_{ij}\lambda_i$  to another state  $j$ , so now, earthquakes occur according to a Poisson process with rate  $\lambda_j$ , or remain in the current state  $i$  with transition rate  $q_{ii}(1) = p_{ii}\lambda_i$ .

Firstly, we need to find the number and the corresponding values of the seismicity rates that are sufficient for the description of the temporal evolution of seismicity. Hence, we fit MAPs from 2 to  $K$  states through the *EM* algorithm to the corresponding trace (inter-event times of the data set). For the determination of the initial values that are given as input to the algorithm, a

grid-based procedure is established implemented in a parallel framework to reduce the computational cost that is increased exponentially with the number of hidden states. All the details on the estimation procedure are given in Section 2.4. Then, the Bayesian or Akaike Information Criterion is used (Equations 2.17 and 2.16, respectively) as a metric for finding the most preferable model, which is the one with the lowest value.

Next, the most probable sequence of transitions for the underlying process,  $J_t$ , is evaluated through the state probabilities,  $p_i(t) = p(J_t = i)$  (Equation 2.14) for  $i = 1, \dots, K$ , whose estimator is a function of the forward and backward equations given in Section 2.6 and its stability is proven on simulated data sets where the true sequence of the hidden states is known (Section 2.6.1). The state of the hidden process at each time  $t$  is defined as the one with the highest probability, i.e.,  $\operatorname{argmax}_{0 \leq i \leq K} p_i(t)$ , with  $p_i(t) = p_i(t_k)$  for  $t_k \leq t < t_{k+1}$ . Each state  $i$  corresponds to an occurrence rate,  $\lambda_i$  (Equation (2.1)), therefore, by evaluating the transitions among the states of the model, we can detect change points in the seismicity rate and the duration of each seismicity level through the sojourn time at each hidden state.

Our main assumption is that due to the multiple embedded occurrence rates of the model each state corresponds to a distinct evolution phase of a seismic sequence, independently of its underlying mechanism. In Section 3.3.1, we evaluate the performance of the model to approximate the temporal evolution of earthquake catalogs that incorporate both main shock-aftershock sequences and earthquake swarms and investigate whether the detected seismicity rate changes are associated with either the occurrence of main shocks and their aftershock sequences or the burst of earthquake swarms.

Concerning the clustering procedure, we are mostly interested in separating periods that are characterized by background and triggered seismicity that is associated with seismic sequences. In particular, we introduce a rate threshold,  $\lambda_{thr}$ , according to which a potential sequence starts when the rate of the counting process,  $N_t$ , achieves  $\lambda_{J_t} > \lambda_{thr}$  and ends as soon as the process  $J_t$

moves for the first time to a state with a Poisson rate below that threshold.

### 3.2.2 Temporal Constraints

The earthquakes that occurred in periods with estimated occurrence rate above the defined rate threshold,  $\lambda_{thr}$ , comprised in the potential clusters. However, results on methods that are solely based on changes in the seismicity rate can sometimes be misleading and additional temporal constraints should be considered. One such case concerns the determination of the aftershock sequences duration. When the rate at the tail of aftershock sequences has reached the level of the background seismicity, it becomes difficult to discriminate these events from background ones, especially for small main shock magnitudes (Felzer and Brodsky, 2006; Godano and Tramelli, 2016). One similar case is related to the sparse foreshock activity, which as it is shown in Lippiello et al. (2019), exhibits significantly smaller frequency than the aftershock activity. Therefore, a day rule,  $dt$ , is assigned in the sense that events in  $\pm dt$  from the potential cluster are included within. Another case is related to the existence of fluctuations during a seismic excitation, when the seismic activity that is triggered by the same underlying mechanism is divided into smaller clusters. A large aftershock sequence can be divided into distinct clusters due to cascade triggering, i.e., main shocks trigger aftershocks which in turn cause their own aftershock sequences (Marsan and Lengline, 2008). For this reason, we assign a time window,  $T$ , so that clusters in temporal distance smaller than or equal to  $T$  are merged into one.

We should note that the choice of the parameters is depending on the goals of each cluster analysis. In a sequence specific study, where the aim is to use the detected clusters for the identification of even the smallest ruptured patches, we would focus on a thinner separation of triggered seismicity (small merging factor,  $T$ ). For studies in large data sets when we intend to investigate the generic clustering properties of seismic sequences, such as the productivity

and temporal decay of aftershocks, the temporal constraints should be looser (larger values for  $dt$  and  $T$ ).

### 3.2.3 DBSCAN algorithm

The grouped events can be spatially sparse and are falsely assigned into the same cluster. To overcome this ambiguity, the DBSCAN is applied to separate events in space based on a distance metric. Depending on the adopted distance metric, the algorithm can be used for grouping events with waveform similarities (Petersen et al., 2021) as well as earthquakes with related types of faulting (focal mechanism similarity) (Cesca, 2020).

Density-based algorithms search for areas where the density level exceeds a threshold. When the spatial density falls below a certain threshold the boundaries are defined and the clusters are formed. The main idea for the DBSCAN algorithm is that for each element of a group of events, the neighborhood of a given radius,  $\epsilon$ , has to include a minimum number of events,  $N_{pts}$ , namely the density has to exceed a threshold value. The geometry of the neighborhood is determined by the adopted distance metric. Hence, the algorithm requires as input the distance threshold,  $\epsilon$ , and the minimum number of neighboring events,  $N_{pts}$  (density level).

The clustering model defines two kinds of events, core points, i.e., events inside the cluster with at least  $N_{pts}$  neighbors within the area including themselves and border points, i.e., events on the border of the cluster. Four main terms are needed to define a cluster:

- Events up to a distance  $\epsilon$  from a core point  $i$  ( $\|N_\epsilon(i)\| \geq N_{pts}$  where  $N_\epsilon(i) = \{j : d(i, j) \leq \epsilon\}$ ) are considered part of the cluster (*directly density reachable*).
- If a neighbor event  $j$  is also a core point then its neighbors are also part of the cluster - the transitive property holds (*density reachable*). In

general, for any two events,  $i_1$  and  $i_n$ , a sequence of directly reachable events,  $i_1, \dots, i_n$ , exists.

- Non-core events are considered border points and they are all density reachable through a common point (*density connected*).
- Events that are not density reachable from a core point are considered *noise* and do not belong to any cluster.

Figure 3.1 shows an example of the DBSCAN method with  $N_{pts} = 3$  and radius  $\epsilon$ . The blue circles correspond to the core points and the arrows indicate the directly density reachable property. Events with yellow color correspond to border points and are density connected since they are both density reachable from any blue event. Finally, the grey event is not density reachable from any other event and is characterised as noise.

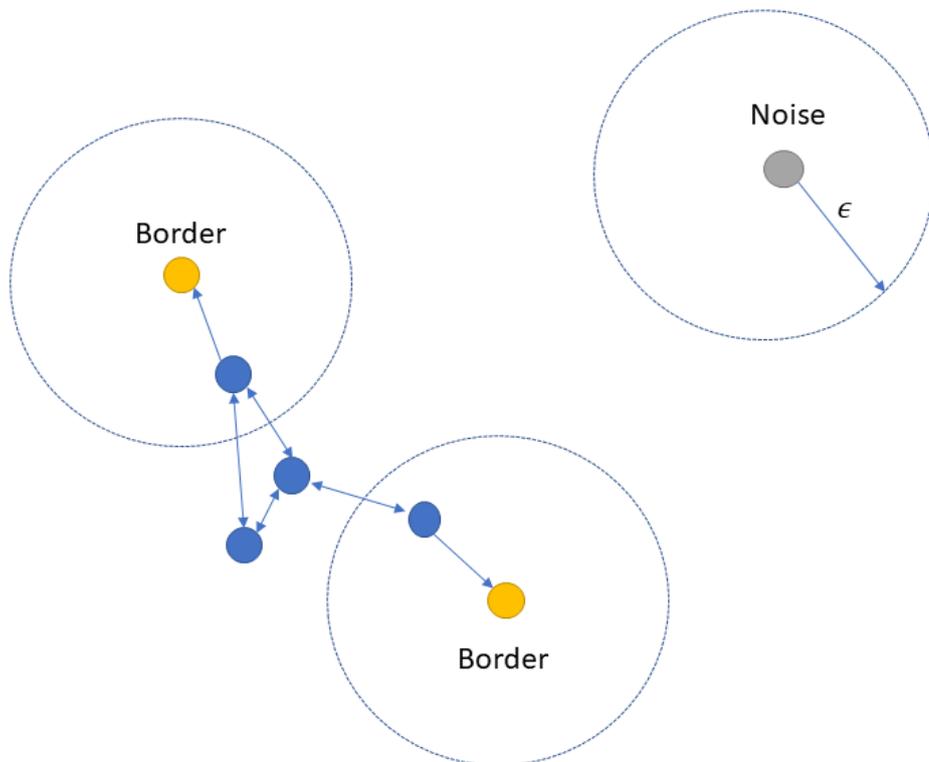


Figure 3.1: An example of the DBSCAN method. Minimum number of neighbors is  $N_{pts} = 3$  and distance threshold equals to  $\epsilon$ .

The algorithm starts with a random event,  $i$ , and forms a cluster if at least  $N_{pts}$  events are within distance  $d \leq \epsilon$ , including itself. Earthquake  $i$  is then considered a core point of the cluster and the algorithm moves to the investigation of the other events. If  $N_{pts}$  neighbors are identified in the next step, they are also considered core events; otherwise, they are the boundary points of the cluster and the algorithm moves to the next unassigned event of the data set.

In our case, events that have not been assigned to any cluster at the end of the procedure compose the background seismicity and are merged with events that occurred during periods with estimated rate under the rate threshold,  $\lambda_{thr}$ . In this way, the algorithm can remove events that are spatially sparse. It has been efficiently applied for detecting similarities among earthquake locations, origin times and focal mechanisms (Bountzidis et al., 2021; Cesca et al., 2016). An advantage of the algorithm is that it does not require as input a pre-defined number of earthquake clusters, such as the k-means algorithm, where further optimization techniques for the determination of the clusters number are necessary (Sheikhhosseini et al., 2021).

### 3.2.4 Evaluation of the MAP-DBSCAN method

The efficiency of the MAP-DBSCAN procedure to correctly identify spatio-temporal correlated seismicity will be evaluated on a simulated ETAS catalog, where the number of clusters and the links among the events are known a priori. The ETAS model is developed by Ogata (1988, 1998) and belongs to a wide class of branching processes where the expected rate of earthquakes, known as intensity function, depends on the history of all previous seismicity. In Appendix A we present an extensional review of the ETAS framework for those not familiar with the model, a thorough description of its parameters and of the simulation procedure. Its intensity function consists of two parts

$$\lambda(t, \mathbf{x} = (x, y)/H_t) = \mu(\mathbf{x}) + K \sum_{j:t_j < t, \mathbf{x}_j \in \Sigma_0} e^{a(m_j - m_c)} g(t) f(\mathbf{x}), \quad (3.1)$$

where  $H_t = \{(t_i, \mathbf{x}_i, m_i) : t_i < t\}$  the subset of all past earthquakes with magnitude larger than or equal to  $m_c$  that occurred inside the region  $\Sigma_0$ . The background component (mother events),  $\mu(\mathbf{x})$ , is stationary in time but heterogeneous in space due to the fault network geometry. For the second term of the right-hand side of Equation 3.1, the modified Omori law (Equation A.3) is considered for the temporal component, the productivity law (Equation A.2) and an isotropic spatial kernel (Equation A.5).

We chose the ETAS model because it captures the main properties of the spatio-temporal evolution of seismicity and it is considered “a de facto standard model, or null hypotheses, for other models and ideas to be compared to” (Huang et al., 2016). It has already been used as a reference model for both forecasting experiments (Nandan et al., 2019) and for evaluating clustering methods such as the Nearest-Neighbor algorithm in Zaliapin and Ben-Zion (2013a).

For the evaluation of the MAP-DBSCAN method we additionally demonstrate its performance against widely used clustering algorithms. In particular, our approach is compared with the Nearest-Neighbor, Gardner and Knopoff window-based and Reasenberg link-based algorithms. A detailed review on each one of them is given in Appendix B.

The simulated ETAS catalog consists of clusters that are defined either as single events, i.e., mother events without any subsequent triggered earthquake or as a sequence of events with a common mother event (first event in the cluster). For the validation of the similarity between the initial structure of the clusters and the one after the implementation of the clustering algorithm, we should first define the true partition of the catalog,  $\mathbf{X} = \{\mathbf{X}_k\}_{k=1, \dots, N_c}$ , where  $N_c$  corresponds to the number of clusters and each subset  $\mathbf{X}_k$  includes



## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

---

all the events of cluster  $k$ . The partition after the implementation of the MAP-DBSCAN and the other  $K - 1$  methods, including different tuning of the algorithm's parameters, is defined as  $\mathbf{Y}_i = \{\mathbf{Y}_n\}_{n=1, \dots, N_{y_i}}$ ,  $i = 1, \dots, K$ , where  $N_{y_i}$  is the number of clusters after the implementation of method  $i$ .

Next, we will use the Jaccard index, which is a measure to quantify the overlap between two partitions, in our case, the true one of the ETAS catalog,  $\mathbf{X}$ , and the one of the  $i$ -th implemented algorithm,  $\mathbf{Y}_i$ . According to [Fortunato and Hric \(2016\)](#) the Jaccard index belongs to the class of similarity measures that are based on pair counting, i.e., counting the number of events assigned to the same (different) cluster. It is expressed by

$$J_1(\mathbf{X}, \mathbf{Y}) = |\mathbf{X} \cap \mathbf{Y}| / |\mathbf{X} \cup \mathbf{Y}| = a_{11} / (a_{11} + a_{10} + a_{01}), \quad (3.2)$$

where  $a_{11}$  indicates the number of pairs of elements which are correctly assigned into the same cluster (true links),  $a_{01}$  the number of pairs of elements which are in the same cluster in the simulated ETAS catalog and in different clusters in the estimated one (missed links) and  $a_{10}$  the number of pairs of elements which are wrongly identified as clustered events (false links). If all the initial clusters are correctly identified by the testing method, then  $a_{10} = 0 = a_{01}$  and  $J_1(\mathbf{X}, \mathbf{Y}) = 1$ . Conversely, if all pairs are wrongly identified as clustered or independent, then  $a_{11} = 0$  and, as a consequence,  $J_1(\mathbf{X}, \mathbf{Y}) = 0$ .

In addition, we introduce a generalization of the Jaccard index,

$$J_2(\mathbf{X}, \mathbf{Y}) = b_{11} / (b_{11} + b_{10} + b_{01}), \quad (3.3)$$

to identify the partition  $\mathbf{Y}$  with the best discrimination between the background seismicity and clustered elements, following the definition in [Lippiello and Bountzlis \(2021\)](#). We consider as background seismicity single events and the mother events of each cluster, i.e., the ones that initiate a cascade of events.  $b_{11}$  represents the number of common background events in the two partitions,



$b_{10}$  is the number of elements wrongly identified as mother events in the partition  $\mathbf{Y}$ , whereas  $b_{01}$  corresponds to the number of true mother events identified as clustered elements in the partition  $\mathbf{Y}$ .

### 3.2.5 Generic ETAS parameters

We will adopt the temporal ETAS model for the investigation of the properties of the detected clusters. It incorporates into its functional form two empirical relationships that characterize the temporal and size distribution of earthquakes, the normalized Omori–Utsu law,  $g(t)$ , given by Equation A.3, and the productivity law,  $N = k(M_i)$  (Equation A.2), where  $N$  is the number of events triggered by an earthquake of magnitude,  $M_i$ .

The constant factor  $K$  in the productivity law is an index of proportionality, which expresses the expected number of triggered events per main shock above the magnitude cutoff, whereas parameter  $a$  describes the impact of magnitude on the number of triggered events. Increasing parameter  $a$  indicates the increase of the proportion of triggered events from large earthquakes compared to small ones typical for main shock–aftershock sequences. Small  $a$  values suggest that the overall productivity is separated into bursts of smaller earthquakes usually the case for earthquake swarms.

To infer the clustering properties of a region we will compute the “generic” ETAS parameters by jointly inverting the ETAS parameter set  $\boldsymbol{\theta} = (p, c, a, K, \mu)$  from the identified clusters. We use the maximum likelihood procedure, which is the most common method for the estimation of the ETAS parameters. In particular, neglecting the spatial component in Equation A.6, the log-likelihood function takes the following form

$$LL_i = \sum_{j=1}^{n_i} \log \lambda(t_j) - \int_{t_0}^{t_f} \lambda(t) dt \quad (3.4)$$

which corresponds to the log-likelihood of the  $i$ -th sequence, namely, the log-



## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

arithmetic probability of observing  $n_i$  events with occurrence times  $t_j$ ,  $j = 1, \dots, n_i$ , during the period of the sequence  $(t_0, t_f)$ . The intensity function,  $\lambda(t)$ , of the model is given by Equation (A.1) and takes the following form when we embed the two empirical laws in the triggering function

$$\lambda(t) = \mu + (p - 1)c^{(p-1)} \sum_{j:t_j < t} K e^{a(M_j - m_c)} \cdot (t - t_j + c)^{-p}. \quad (3.5)$$

Then, we stack all the detected sequences of the study area, compute their corresponding logarithmic probabilities,  $LL_i$ , and define as the common log-likelihood

$$LL = \sum_{i=1}^{N^*} LL_i, \quad (3.6)$$

where  $N^*$  is the number of sequences. The optimal inverted parameters are the ones that maximize Equation (3.6). For the maximization of the common log-likelihood,  $LL$ , we implement the numerical procedure in A.2. However, in this case at each iteration step,  $(r)$ , we compute  $N^*$  log-likelihood values,  $LL_i^{(r)}$ ,  $i = 1, \dots, N^*$ , and store the new parameters under the condition  $LL^{(r)} > LL^{(r-1)}$  where  $LL$  is the common log-likelihood given by Equation (3.6).

For the comparison of the aftershock productivity among areas with different completeness magnitudes, we adopt a common magnitude cutoff (the maximum) and use the following relation,

$$N = k(M_i)P(M \geq m_c^*) = K e^{a(M_i - m_c)} e^{-\beta(m_c^* - m_c)}, \quad (3.7)$$

which yields the number of earthquakes above magnitude  $m_c^*$ , generated by a main shock of magnitude  $M_i$ . The exponent of the exponential magnitude distribution is expressed by  $\beta$  and is defined as  $\beta = \sum_{i=1}^{N^*} \beta_i / N^*$ , where  $N^*$  is the number of clusters for each study area and  $\beta_i$  their corresponding exponent values.

### 3.3 Applications

We firstly apply the MAP model to the Corinth Gulf area and evaluate its efficiency to separate potential seismic sequences from background seismicity in an earthquake catalog. In the second part, we will apply the two-step clustering procedure, MAP-DBSCAN, first on a synthetic earthquake catalog where the structure of the clusters is known a-priori and subsequently to three major seismic zones of Greece with different seismotectonic properties. The selected areas are the Corinth Gulf, the Central Ionian Islands and the North Aegean Sea, all characterized by intense seismic activity and therefore a large amount of data is available. We will investigate their regional clustering properties based on the ETAS model and will produce generic and sequence specific parameters for each study area.

#### 3.3.1 Seismicity rate changes through MAP in Corinth Gulf, Greece

In this section we will test the efficiency of the MAP model to detect seismicity rate changes in Corinth Gulf area. The MAP is a temporal model that does not take into account spatial correlations, so a sufficiently small region for its application is assumed for any given earthquake to may interact with all following events, regardless of their spatial locations, as suggested by [Zhuang et al. \(2012\)](#). The Corinth Gulf area is divided into its eastern and western parts, based on seismotectonic criteria, like fault segmentation, slip rates and variations in seismicity rates ([Bountzidis et al., 2019](#); [Console et al., 2013](#)). We note, that a number of strong ( $M \geq 6.0$ ) events outside the study area could have a triggering effect, which could further improve the fitting of the model, however, due to lack of a spatial component we are extremely reluctant into extending the area.

The 1964-2017 earthquake catalog was used and is considered complete for earthquakes with  $M \geq 4.5$  (Console et al., 2015) including 274 events within an area bounded between  $21.3^\circ - 23.2^\circ$  and  $37.9^\circ - 38.6^\circ$ . The high magnitude cutoff enables the use of a long period earthquake catalog that includes well studied swarm type sequences as well as an adequate number of strong main shocks. This will allow us to evaluate the ability of the model on capturing changes in the seismicity rate related to both main shock-aftershock sequences and earthquake swarms. In Table 3.1 details on the sub-catalogs of the eastern and western subareas are given and Figure 3.2 shows with yellow stars the epicenters of the 274 events with  $M \geq 4.5$  during 1964-2017.

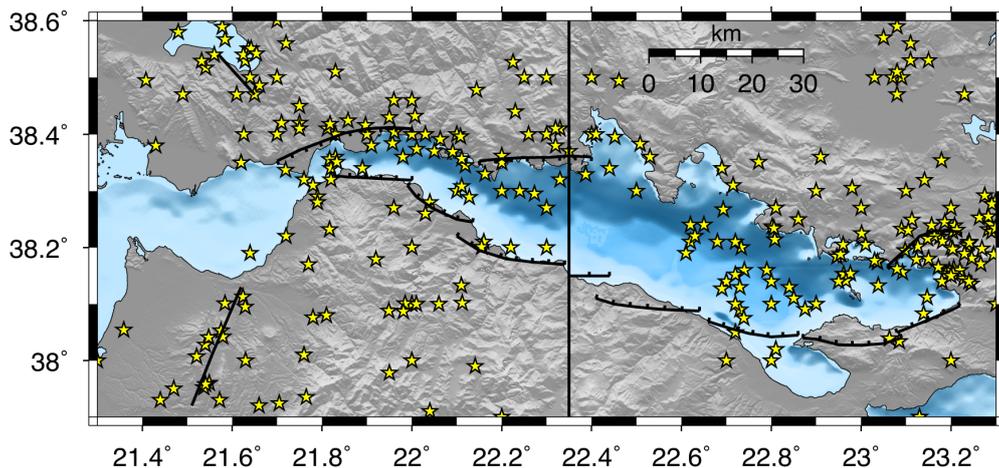
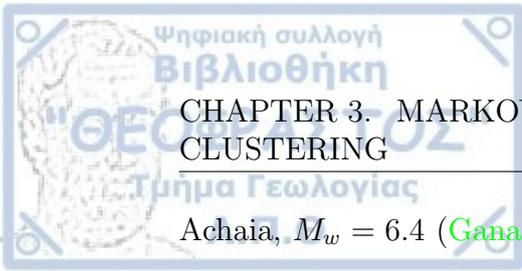


Figure 3.2: Epicentral distribution of earthquakes with  $M \geq 4.5$  that occurred in the area of Corinth Gulf during 1964-2017 are shown with yellow stars. Vertical line divides the area into the western and eastern subareas.

Among the seismic excitations with special interest in the eastern subarea is the sequence of three strong earthquakes ( $M \geq 6.3$ ) that occurred during nine (9) days in February-March 1981, associated with adjacent and antithetic fault segments in eastern Corinth Gulf (Papazachos et al., 1984). In the western subarea, recent strong earthquakes include the 1993 Patras,  $M_w = 5.6$  (Karakostas et al., 1994), 1995 Aigion,  $M_w = 6.4$  (Bernard et al., 1997), 2008



## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Achaia,  $M_w = 6.4$  (Ganas et al., 2009; Karakostas et al., 2017), and the 2010 Efpalio doublet, with  $M_w = 5.5$  and  $M_w = 5.4$  (Karakostas et al., 2012). Table 3.1 provides information on these strong earthquakes.

Table 3.1: Strong earthquakes ( $M \geq 5.5$ ) in Corinth Gulf area during 1964-2017.  $T$  and  $N$  denote the study period and the number of events comprised in each data set, respectively.

Date	Time	Lat.	Lon.	$M_w$		$M_{thr}$	$T$	$N$
1965, Jul 6	03:18:42	38.270	22.300	6.3	Corinth Gulf	4.5	1964-2017	274
1975, Jun 30	13:26:55	38.466	21.641	5.6				
1975, Dec 31	09:45:45	38.486	21.661	6.0	Western subarea	4.5	1964-2017	139
1981, Feb 24	20:53:37	38.153	22.961	6.7				
1981, Feb 25	02:35:51	38.083	23.139	6.4	Eastern subarea	4.5	1964-2017	135
1981, Mar 4	21:58:05	38.204	23.236	6.3				
1993, Jul 14	12:31:49	38.170	21.770	5.6				
1995, Jun 15	00:15:50	38.362	22.200	6.4				
2008, Jun 8	12:25:28	37.952	21.537	6.4				
2010, Jan 18	15:56:9	38.404	21.961	5.5				

### 3.3.1.2 Model evaluation and comparison with MMPP

Firstly, we need to determine the optimal number of states for the MAP model. Thus, MAPs with two to four states are fitted to the events with  $M \geq 4.5$  that occurred during 1964-2017 in both subareas. We believe that a MAP with four states is sufficient as an upper limit of state space for the description of the earthquakes temporal distribution (Bountzlis et al., 2019). The AIC and BIC values of the fitted models as well as the corresponding log-likelihood values are given in Table 3.2. Both criteria are in favor of the four-states model for the eastern subarea (lowest values with AIC=916 and BIC=974), however, in the western subarea AIC and BIC suggest the four and three-states model (AIC=1395 and BIC=1443), respectively.

CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Table 3.2: The log-likelihood and AIC and BIC values of the fitted MAP models to the events with  $M \geq 4.5$  that occurred in the eastern and western subarea of Corinth Gulf during 1964-2017.

# of S	Western subarea			Eastern subarea		
	$LL$	AIC	BIC	$LL$	AIC	BIC
2	-708	1429	1447	-496	1005	1023
3	-692	1408	1443	-461	946	981
4	-677	1395	1453	-438	916	974

For each MAP, the transformed times are calculated with the relation (2.18) and the cumulative number of the residuals versus the transformed times are plotted in Figure 3.3 along with the Kolmogorov-Smirnov (KS) confidence bounds. A good approximation of the true process is achieved when the residuals follow a stationary Poisson process with unit rate (solid line). For the three-states MAP in the western subarea (Figure 3.3a) the 95% confidence bounds are almost exceeded in three cases, whereas for the four-states MAP the residual process is well approximated (Figure 3.3b). Similar performance with the latter shows the four-states MAP of the eastern subarea (Figure 3.3c).

Concerning the inter-event times,  $E_i = \tau_{i+1} - \tau_i$ ,  $i = 1, \dots, n - 1$ , with  $n$  being the number of events, two types of tests need to be implemented for the assumption of a stationary Poisson process to hold. Firstly, the Runs test and the auto-correlation function are computed in order to seek for possible correlations and next Kolmogorov-Smirnov test is applied to verify that the  $E_i$  follow an Exponential distribution. Figure 3.4 shows the auto-correlation of  $E_i$  and Table 3.3 gives the  $p$ -values of the Runs and KS tests for each of the three fitted MAP models. According to the  $p$ -values of the Runs test, we cannot reject the null hypothesis of uncorrelated events ( $p > 0.05$ ), which is also verified from the auto-correlation plot (all values within the confidence bounds). Concerning the KS goodness-of-fit test, besides the three-states MAP of the western subarea of Corinth Gulf where  $p = 0.009$ , there is no evidence for a significant departure from the stationary Poisson process hypothesis. Therefore, we choose the four-states MAP as the optimal model that captures

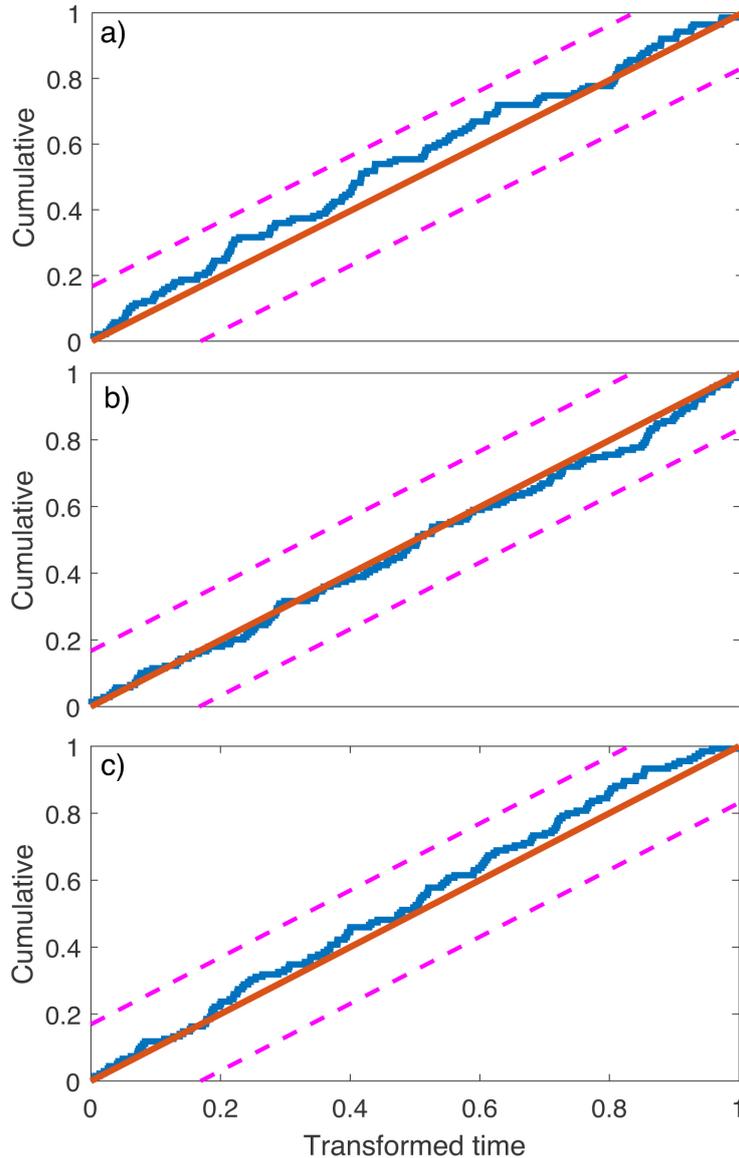


Figure 3.3: The residual process for the western subarea with the a) three-states MAP and b) four-states MAP and for the eastern subarea with the c) four-states MAP. The solid straight line corresponds to the stationary Poisson process with unit rate and the dashed lines to the 95% Kolmogorov-Smirnov confidence bounds.

the main temporal features of seismicity with  $M \geq 4.5$  in both subareas.

Table 3.4 presents the estimated parameters of the rate matrices,  $\mathbf{D}_0$  and  $\mathbf{D}_1$  for the four-states MAPs of the western and eastern subarea of Corinth Gulf, respectively. We should note that a MAP model with an extra hidden state (corresponds to occurrence rate,  $\lambda_4^{west} = 182.8261$ ) is found compared

CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Table 3.3: P-values of the Runs and Kolmogorov-Smirnov tests implemented to the inter-event times of the residuals,  $E_i$ , for the three-states and four-states MAP of the western subarea of Corinth Gulf and for the four-states MAP of the eastern subarea of Corinth Gulf.

Statistical tests	Western subarea		Eastern subarea
	MAP(3)	MAP(4)	MAP(4)
Runs test	0.775	0.536	0.247
Kolmogorov-Smirnov test	0.009	0.199	0.999

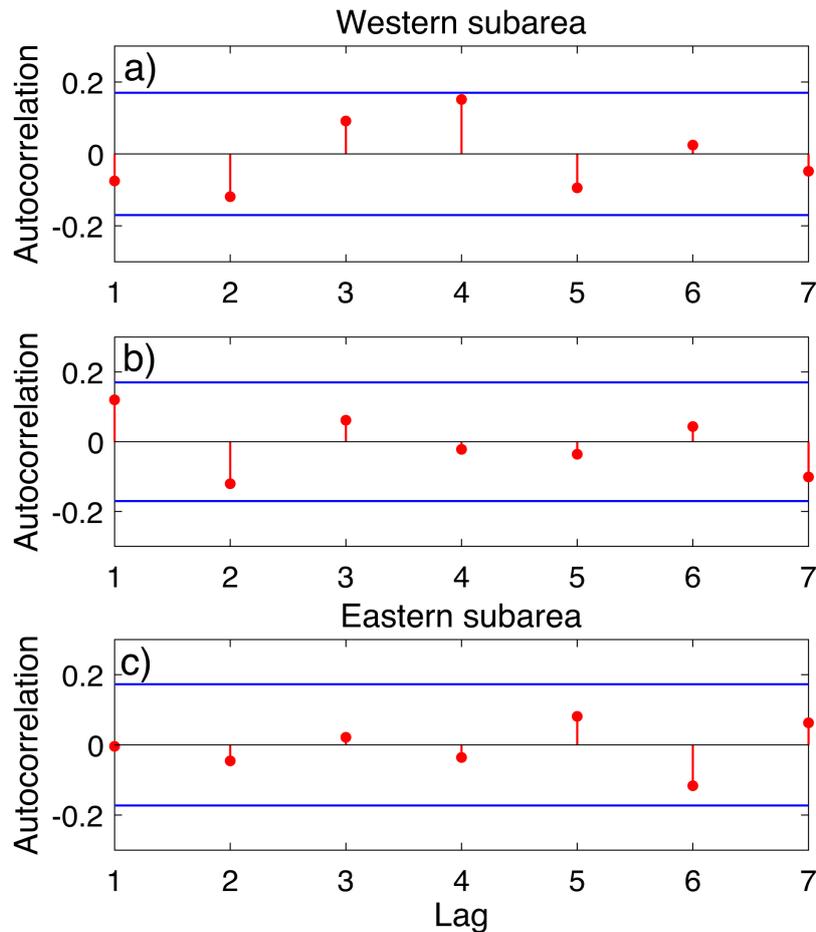


Figure 3.4: The auto-correlation of the inter-event times,  $E_i$ , for the western subarea of Corinth Gulf with the a) three-states MAP and b) four-states MAP and c) for the eastern subarea of Corinth Gulf with the four-states MAP. Solid blue lines indicate the 95% confidence bounds.

to [Bountzlis et al. \(2019\)](#) where the same MAP model is applied to the western subarea for earthquakes with  $M \geq 4.5$  during 1964-2016. The use of the grid-based procedure for the choice of the initial input values of the  $EM$  algo-

## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

gorithm, led to a higher peak of the log-likelihood for the four-states MAP, large enough to be in favor over the three-states model. The investigation of the log-likelihood landscape profile shows an increased number of peaks as the number of states is getting higher, which makes the choice of the initial parameter set really crucial. Therefore, the condition under which the *EM* algorithm was implemented in the previous work, when initial parameter set was chosen to match the first moment of the trace, seems very sensitive when small data sets are combined with large parameter sets, which is the case here. Nevertheless, characteristics concerning the dynamics of the two subareas do not differ significantly, as the episodic seismicity behavior of the eastern subarea compared to the western one still remains. According to the estimated transition rates given in Table 3.4 the probability to move from the state with the highest seismicity rate (state 4,  $\lambda_4^{west} = 182.82$ ,  $\lambda_4^{east} = 41.72$  events/day), to the one with the lowest occurrence rate (state 1,  $\lambda_1^{west} = 0.0022$ ,  $\lambda_4^{east} = 0.0029$  events/day), is higher in the eastern ( $p_{41}^{east} = q_{41}(1)/q_{44}(0) = 0.1054$ ) than in the western subarea ( $p_{41}^{west} = 4.0516e - 34$ ) and vice versa ( $p_{14}^{east} = 0.0716$ ,  $p_{14}^{west} = 2.2459e - 07$ ). This is an indicator of a more bursty behavior for the eastern subarea, i.e., faster transitions from periods with high productivity to periods of relative seismic quiescence.

Table 3.4: Parameter estimates of the  $\mathbf{D}_0$  and  $\mathbf{D}_1$  matrices corresponding to the four-states MAPs for the events with  $M \geq 4.5$  that occurred in western and eastern subarea of Corinth Gulf during 1964-2017. Time unit is in days.

Subarea	Estimated Parameters								
	(0,1)	(0,2)	(0,3)	(0,4)	(1,1)	(1,2)	(1,3)	(1,4)	
Western	(0,1)	-0.0022	0	0	0	2.9021e-07	0.0017	0.0004	4.8447e-10
	(0,2)	0	-0.0113	0	0	0.0036	0.0065	0.0003	0.0007
	(0,3)	0	0	-2.6586	0	1.6734e-10	1.3771	1.1527	0.1287
	(0,4)	0	0	0	-182.8261	7.4073e-32	44.1805	68.9208	69.7246
Eastern	(0,1)	-0.0029	0	0	0	0.0026	6.4091e-05	8.8485e-11	0.0002
	(0,2)	0	-0.0859	0	0	0.0058	0.0656	0.0082	0.0061
	(0,3)	0	0	-2.6525	0	0.0844	0.2644	2.1382	0.1653
	(0,4)	0	0	0	-41.7287	4.4002	3.4708e-08	6.3211	31.0073

We also moved to the fitting of a MMPP on the sub-catalogs of the western and eastern subareas in order to compare it with the corresponding fitted MAP models. First,  $N = 1E6$  initial values are randomly chosen and the parameter

set corresponding to the maximum log-likelihood is selected as input value to the *EM* algorithm. The Hidden Markov package developed by [Harte \(2017\)](#) is used and the convergent log-likelihood values for the eastern ( $LL = -456$ ) and western ( $LL = -690$ ) subareas are given in Table 3.5, respectively. In addition, the AIC and BIC differences between the two models ( $\Delta AIC = AIC_{MMPP} - AIC_{MAP}$ ,  $\Delta BIC = BIC_{MMPP} - BIC_{MAP}$ ) are computed for both subareas. Positive values indicate that the MAP model is preferred compared to the MMPP and vice versa. In both data sets the differences (Table 3.5) are positive which means that with the MAP model we achieve a better fitting.

Table 3.5: The Akaike's and Bayes Information Criteria differences among the Markov Modulated Poisson Process and the Markovian Arrival Process for the western and eastern subareas with  $M \geq 4.5$  during 1964-2017 and the corresponding log-likelihood values for the MMPP.

Area	$LL_{MMPP(4)}$	$LL_{MAP(4)}$	$\Delta AIC$	$\Delta BIC$
Eastern subarea	-456	-438	28.2531	16.6618
Western subarea	-690	-677	18.4433	6.7343

### 3.3.1.3 Identified seismic sequences

Figure 3.5a visualizes the transitions among the states of the four-states MAP for the data set of the western subarea of Corinth Gulf with the use of colored boxes. The color at each temporal interval  $t_k \leq t < t_{k+1}$  indicates the state with the maximum probability at the certain time, derived with Equation 2.14, and the legend contains its corresponding occurrence rate, given by Equation 2.1. Red color indicates the sojourn of the process in state one ( $\lambda_1 = 0.002$ ), yellow to state two ( $\lambda_2 = 0.01$ ), orange to state three ( $\lambda_3 = 2.66$ ) and cyan to state four ( $\lambda_4 = 182.83$ ). The intensity function of the MAP model is also evaluated by Equation 2.15 and expresses the expected occurrence rate,  $\lambda(t) = \lambda(t_k)$ ,  $t_k \leq t < t_{k+1}$  at each time  $t$ . This is an alternative way to illustrate the fluctuations in the seismicity rate during the study period.

We observe more frequent fluctuations of the intensity function (Figure

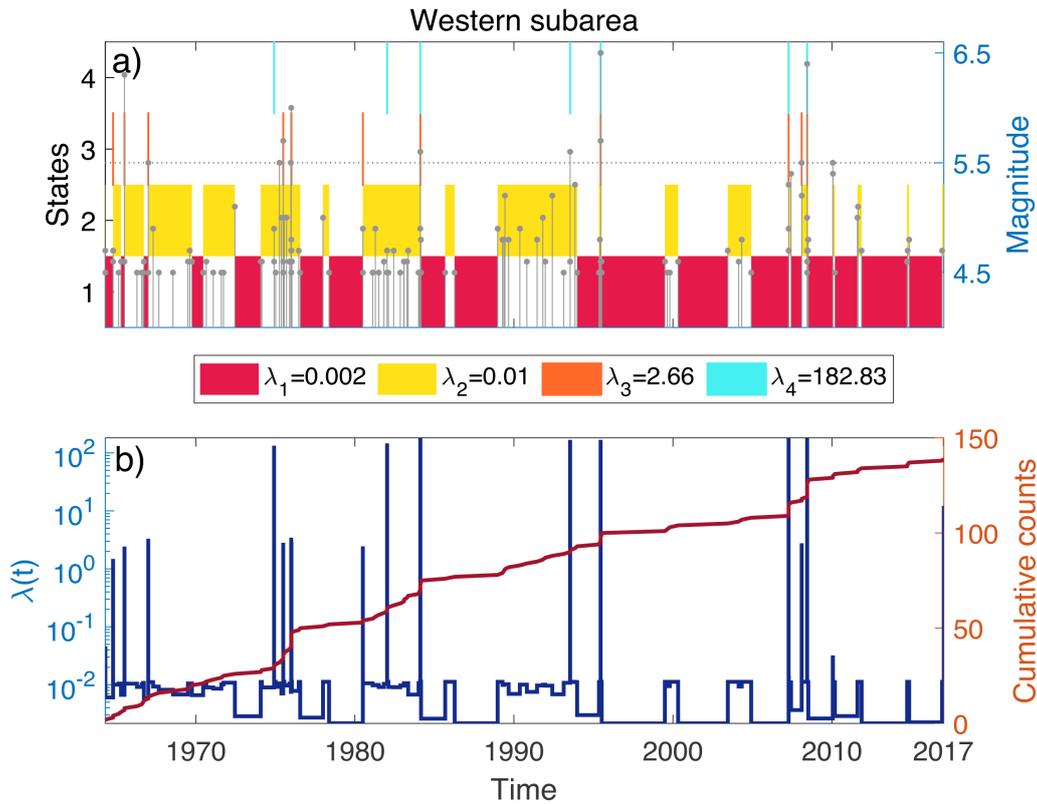


Figure 3.5: a) The most probable sequence of hidden states along with the magnitude distribution as a function of time (grey vertical lines) for the events with  $M \geq 4.5$  of the western subarea of Corinth Gulf during 1964-2017. Each color is assigned to a different state  $i$  with seismicity rate  $\lambda_i$ . Red color corresponds to  $\lambda_1 = 0.002$ , yellow color to  $\lambda_2 = 0.01$ , orange color to  $\lambda_3 = 2.66$  and cyan color to  $\lambda_4 = 182.83$ . (b) The expected seismicity rate is denoted by the blue color and the cumulative number of the events with the maroon.

3.5b) until 1995, where the majority of strong earthquakes with  $M \geq 5.5$  occurred. This can also be seen in Figure 3.5a from the more frequent visits of the model to states 3 and 4, which are associated with the highest seismicity rates ( $\lambda_3 = 2.66$ ,  $\lambda_4 = 182.83$ ). After 1995 we see that state 1 with the lowest occurrence rate ( $\lambda_1 = 0.002$ ) is dominant compared with the previous period where the sojourn time of the model to state 2 ( $\lambda_2 = 0.01$ ) is longer. States 1 and 2 can be associated to periods of relative seismic quiescence with  $\lambda_1 = 0.002$  and  $\lambda_2 = 0.01$ , respectively, and separate the data from periods of high (states three and four) seismic activity that might be related with a

## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

seismic excitation. For this reason, we focused on the events that occurred during the sojourn of the process in states 3 and 4. We observe that strong earthquakes ( $M \geq 5.5$ ) with special interest mentioned in Table 3.1 and their associated events are related to these states.

Table 3.6: Information on the earthquakes that occurred during the sojourn of the process in states three and four in the western subarea of the Corinth Gulf during 1964-2017. # denotes the potential cluster index, second column corresponds to the time interval, "No" denotes the number of events occurred during the sojourn of the process in states three and four, in the fourth column the main shocks are presented and in the last column some details are given.

#	Time interval	No	Main shocks	Details
1	6/07/1965	2	$M = 6.3$ (state 3)	The main shock and its aftershock just 3 hours later occurred in state 3 (Figure 3.6a).
2	4/01/1967	2	$M = 5.5$ (state 3)	The main shock and an aftershock is observed after one hour in state 3 (Figure 3.6b).
3	29-30/06/1975	3	$M = 5.7$ (state 3)	A foreshock one day before the 30 Jun 1975 main shock with $M = 5.7$ is observed, and one aftershock in 5 hours. All events occurred in state 3 (Figure 3.6c).
4	31/12/1975-3/01/1976	7	$M = 6.0$ (state 3)	The 31 December 1975 main shock of $M = 6.0$ occurred, and an aftershock sequence followed with 6 events in four days (Figure 3.6d).
5	11/12/1984	3	$M = 5.6$ (state 4)	The model captures the main shock-aftershock sequence (Figure 3.6e). However, the two events that occurred close in space during the last month before the main shock, and the secondary aftershock sequence of two more events during the next 17 days occurred in state 2, probably associated with the relative increase and decrease in the seismicity before and after the main shock, respectively.
6	14/07/1993	2	$M = 5.6$ (state 4)	The July 14, 1993 ( $M = 5.6$ ) Patras earthquake (Karakostas et al., 1994) and the immediate aftershock with $M = 4.6$ (eight minutes later) (Figure 3.6e).
7	15/6/1995	3	$M = 6.5$ (state 4)	The 1995 Aigion sequence (Bernard et al., 1997) is revealed. The $M = 6.5$ , June 15, 1995 Aigion earthquake together with two immediate aftershocks (in almost 4 hours after the main shock) occurred in state 3, whereas two foreshocks and two more aftershocks (state 2) occurred one month before and after the main shock, respectively. (Figure 3.6g).
8	9-10/04/2007	6	$M = 5.3$ (state 3)	The April 2007 earthquake swarm occurred near the Lake Trichonis (Kıratzi et al., 2008). An intense seismic sequence is observed where three strong earthquakes, all of magnitude $M = 5.3$ , along with three events of magnitude $4.5 \leq M \leq 4.9$ that occurred in less than 24 hours, all in states 3 and 4. (Figure 3.6h).
9	4/02/2008	2	$M = 5.5$ (state 3)	A main shock with $M = 5.5$ and its immediate aftershock with $M = 5.2$ (in less than two hours) occurred in state 3 (Figure 3.6i).
10	8-9/06/2008	8	$M = 6.4$ (state 4)	The 2008 Achaia sequence (Ganas et al., 2009; Karakostas et al., 2017) is revealed. The $M = 6.4$ June 12, 2008 main shock together with four immediate aftershocks (within an hour after the main shock) occurred in state 4 moving afterwards to state 3 where three aftershocks occurred until next day. (Figure 3.6j).

Table 3.6 details the events that occurred during the periods of high seismicity rate according to the model ( $\lambda_3 = 2.66$ ,  $\lambda_4 = 182.83$ ) and Figure 3.6

shows their epicentral distribution. State 3 and 4 correspond in most cases to main shocks and their immediate aftershocks. For instance, the intense aftershock sequence of December 1975 (cluster 4 in Table 3.6, Figure 3.6d) close to Lake Trichonis (Kiratzi et al., 2008) was recorded during the sojourn of the model in state 3. In 31 December 1975 a main shock of  $M = 6.0$  occurred, that was followed by an aftershock sequence of 6 events in four days. In cluster 7 (Table 3.6), the main shock ( $M = 6.5$ ) is associated to state four, which is followed by an aftershock 15 minutes later moving to state three (Figure 3.6g). The April 2007 earthquake swarm that occurred near the Lake Trichonis (Kiratzi et al., 2008) is also associated with the sojourn of the process in states 3 and 4. An intense seismic sequence is observed with three strong earthquakes, all of magnitude  $M = 5.3$ , along with three events of magnitude  $4.5 \leq M \leq 4.9$  all within 24 hours. The epicenters of the swarm are in close proximity to the epicenters of the June 1975 (cluster 3 in Table 3.6) and December 1975 (cluster 4 in Table 3.6) sequences, both unveiled by the underlying process (Figures 3.6c,d,h).

In general, we conclude that transitions to or from states 3 and 4 indicate the initiation or the end of a seismic sequence with  $M \geq 4.5$ , respectively. There are only two groups of events that lack spatial proximity and occurred in states 3 and 4, which cannot be associated to a known seismic sequence. A sojourn period in state 2 without a transition to states with higher seismicity rates (states 3 and 4), might be related to an earthquake swarm or a main shock sequence where the maximum magnitude is close to the completeness threshold,  $M_c = 4.5$ . In this case, events with  $M < 4.5$  that can be part of the sequence, are not included in the data set and therefore they were not considered when we estimated the occurrence rates. This is the case of the two moderate events ( $M = 5.5$ ,  $M = 5.4$ ) that occurred in January 2010 during the sojourn of the process in state 2, where the main seismic activity consists of events with  $M < 4.5$  (Karakostas et al., 2012). Finally, all eight (8) events of the data set with  $M \geq 5.6$  occurred in state 3 or 4, and since

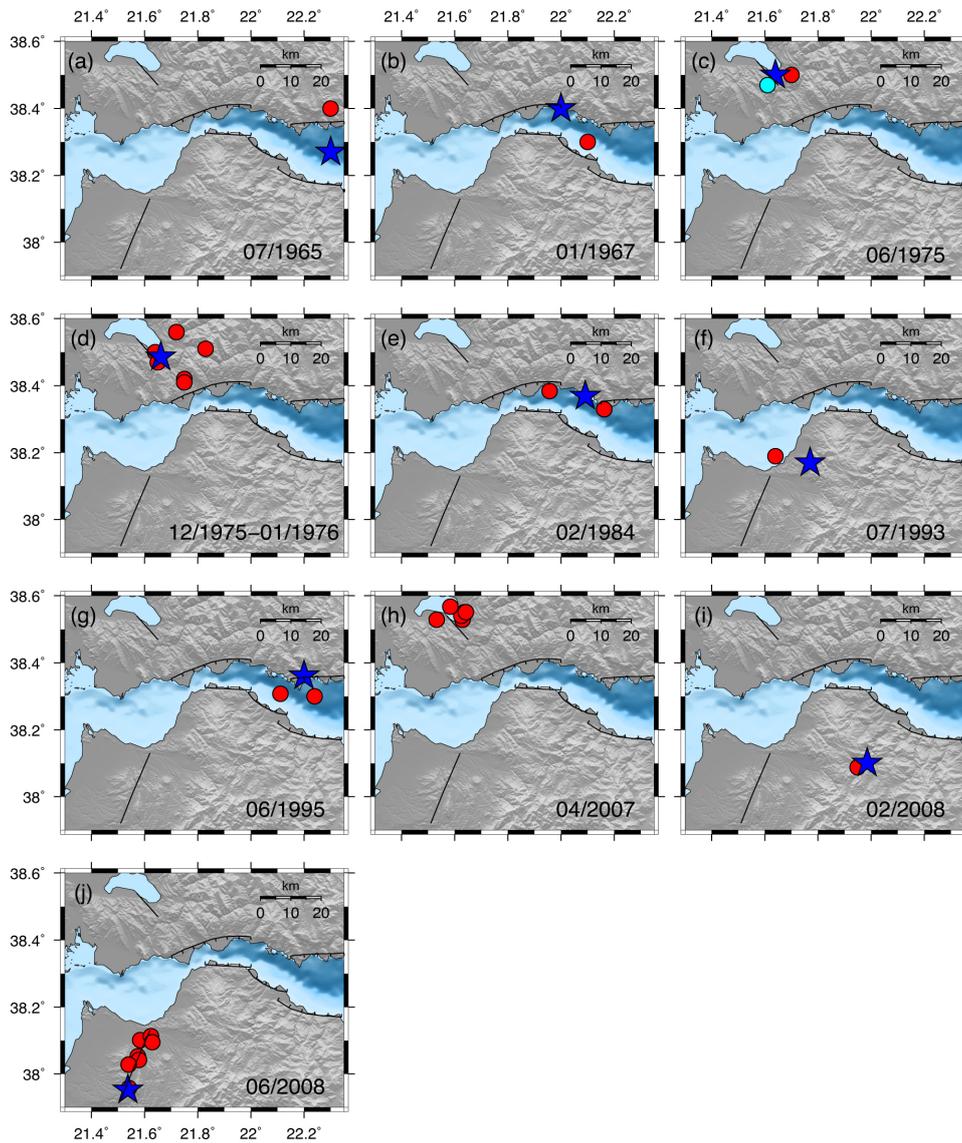


Figure 3.6: Epicentral distribution of the earthquakes that occurred during the sojourn of the process in states 3 and 4 in the western subarea during 1964-2017. Stars (blue) represent main shocks and circles correspond to earthquakes that occurred before (cyan) and after (red) the main shock, respectively. Earthquakes associated with the (a) 6 July 1965, (b) 4 January 1967, (c) 30 June 1975, (d) 31 December 1975 (e) 11 February 1984 main shocks. (f) The main shock of the 14 July 1993. (g) The Aigion sequence on June 1995. (h) The April 2007 earthquake swarm near Lake Trichonis. (i) Earthquakes associated with the 4 February 2008 and (j) 8 June 2008 main shocks.

the mean sojourn time until the next earthquake occurrence is  $1/\lambda_3 = 9$  hours and  $1/\lambda_4 = 8$  minutes, respectively, we may conclude that an event with

magnitude  $M \geq 4.5$  is expected in the next few hours after states 3 or 4 are reached, namely after the occurrence of an event with  $M \geq 5.6$ .

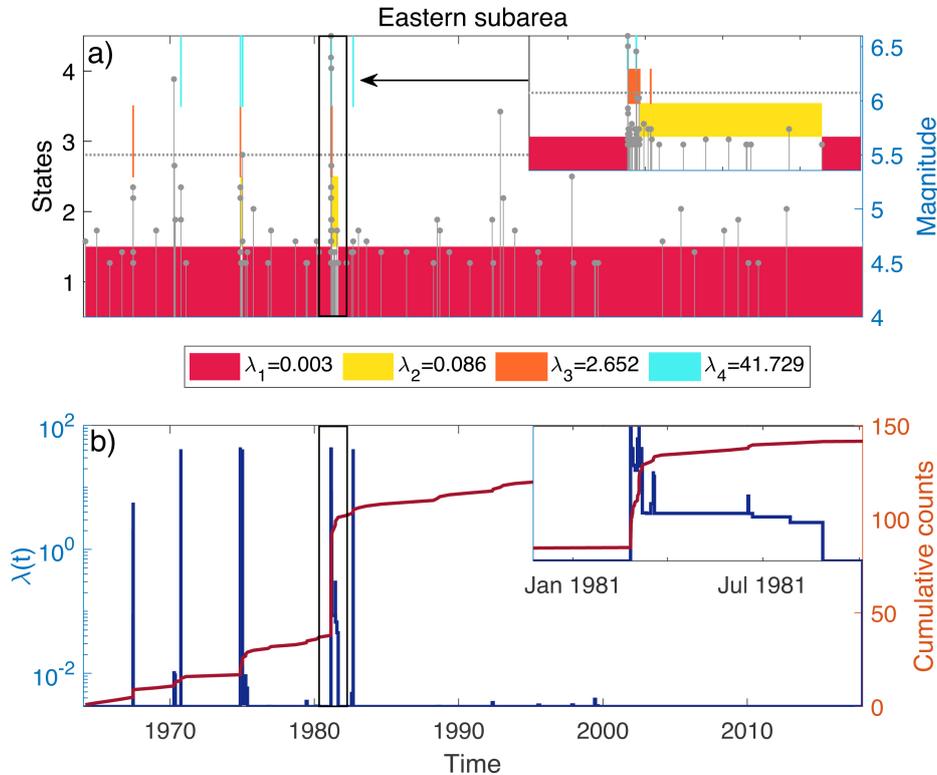


Figure 3.7: a) The most probable sequence of hidden states along with the magnitude distribution as a function of time (grey vertical lines) for the events with  $M \geq 4.5$  of the eastern subarea of Corinth Gulf during 1964-2017. Each color is assigned to a different state  $i$  with seismicity rate  $\lambda_i$ . Red color corresponds to  $\lambda_1 = 0.002$ , yellow color to  $\lambda_2 = 0.01$ , orange color to  $\lambda_3 = 2.66$  and cyan color to  $\lambda_4 = 182.83$ . (b) The expected seismicity rate is denoted by the blue color and the cumulative number of the events with the maroon. Magnified in the inset is the evolution of the Alkyonides seismic sequence according to the hidden path of the model.

Figure 3.7 presents the most probable path of the hidden states along with the intensity function for the four-states MAP of the eastern subarea of Corinth Gulf, for earthquakes with  $M \geq 4.5$  during 1964-2017. As before, red colored boxes represent the sojourn of the underlying process to state one ( $\lambda_1 = 0.003$ ), yellow boxes to state two ( $\lambda_2 = 0.086$ ), orange boxes to state three ( $\lambda_3 = 2.652$ ) and cyan boxes to state four ( $\lambda_4 = 41.729$ ). In western subarea 13 events with  $M \geq 5.5$  occurred, whereas in the eastern subarea these are only 6 events

## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

with  $M \geq 5.5$ . Consequently, according to the hidden path longer quiescent periods are observed (state 1 with  $\lambda_1 = 0.003$ , red color in Figure 3.7a), which are interrupted by intense seismic periods, i.e., transitions to states with higher seismicity rates (states 2, 3 and 4 with  $\lambda_2 = 0.086$ ,  $\lambda_3 = 2.652$  and  $\lambda_4 = 41.729$ , respectively, in Figure 3.7a). Large variations of the seismicity rate are also observed from the model's intensity function in Figure 3.7b. Compared with the corresponding plot for the western subarea (Figure 3.5b), the expected occurrence rate is more stable, especially since 1982. The productivity is very low with state 1 ( $\lambda_1 = 0.003$ ) being dominant during the whole study period. State 1 is associated to periods of relative seismic quiescence, and we can say that it represents the background rate for the data set. States 2, 3 and 4 are related to bursts of activity, so we focus on events that occurred during the sojourn of the model in these states.

Table 3.7: Information on the earthquakes that occurred during the sojourn of the process in states 2, 3 and 4 in the eastern subarea of Corinth Gulf during 1964-2017. # denotes the potential cluster index, second column corresponds to the time interval, "No" denotes the number of events occurred during the sojourn of the process in these states, in the fourth column the main shocks are presented and in the last column some details are given.

#	Time interval	No	Main shocks	Details
1	8/6/1967-12/6/1967	5	$M = 5.2$ (state 2)	The 8 Jun 1967 main shock ( $M = 5.2$ ) occurred in state 2 and four aftershocks followed in the next 4 days (Figure 3.8a).
2	14/11/1974-4/12/1974	6	$M = 5.2$ (state 4)	A foreshock and two aftershocks occurred in less than a day after the main shock ( $M = 5.2$ ). Then the model moved to state 2 ( $q_2 = 0.0859$ ) until the beginning of the seismic activity in the adjacent area (Figure 3.8b).
3	1/1/1975-8/1/1975	4	$M = 5.5$ (state 4)	Two events before the main shock ( $M = 5.5$ ) occurred and one aftershock within one hour (Figure 3.8c).
4	24/02/1981-27/08/1981	64	$M = 6.6, 6.4, 6.3$ (state 4)	The February-March, 1981 seismic sequence is remarkably well revealed through the path of the hidden states (Figure 3.7). The major part of the February 24th main shock aftershock sequence occurred during the sojourn time in state 4 (14 events in approximately 10 hours). In the sequel, another transition to state 3 occurred not long before the occurrence time of the third destructive event on March 4, in the northeastern part of the gulf. In 8 Mar the model moved to state 2 up until 27 Aug (Figure 3.8d).

The five groups of events (consecutive events in states 2, 3 and 4) corre-

spond to potential clusters (concentrated in space and time) and four of them are detailed in Table 3.7. Their epicentral distribution is shown in Figure 3.8. We should mention, however, that clusters 2 and 3 were recognized as one group, probably due to their temporal proximity. The underlying process with the four hidden states succeeded to reveal the three moderate main shocks along with their aftershocks, the June 1967 (Figure 3.8a), the November 1974 (Figure 3.8b) and the January 1975 (Figure 3.8c) sequences as well as the complicated temporal structure of the February-March, 1981 Alkyonides seismic sequence (Figure 3.8d), validating previous observations.

To sum up, the multiple embedded occurrence rates of the underlying Markov process  $J_t$  capture the seismicity temporal evolution of earthquakes with  $M \geq 4.5$  in Corinth Gulf during 1964-2017, and reveal the seismicity rate changes associated in many cases with the occurrence of main shocks and their aftershock sequences or earthquake swarms.

### 3.3.2 Application of MAP-DBSCAN

In this section, we firstly test the ability of the two-step clustering procedure to detect spatiotemporal clusters on an simulated ETAS earthquake catalog where the number of clusters and the links among the events are known a priori.

Subsequently, we apply the procedure to earthquake data sets and study the properties of the detected clusters. For the application of the MAP-DBSCAN procedure, we firstly need to define the earthquake catalog, specifically, the spatial boundaries, the time window and the completeness magnitude. The former is crucial since the structure of the clusters is related to the seismotectonic properties of a region. Seismic sequences can extend outside the boundaries if they are chosen arbitrarily, which produces misleading results (Wang et al., 2010). Hence, it is important to define the boundaries of the study area, taking into account the tectonic setting and the seismic activity. The time win-

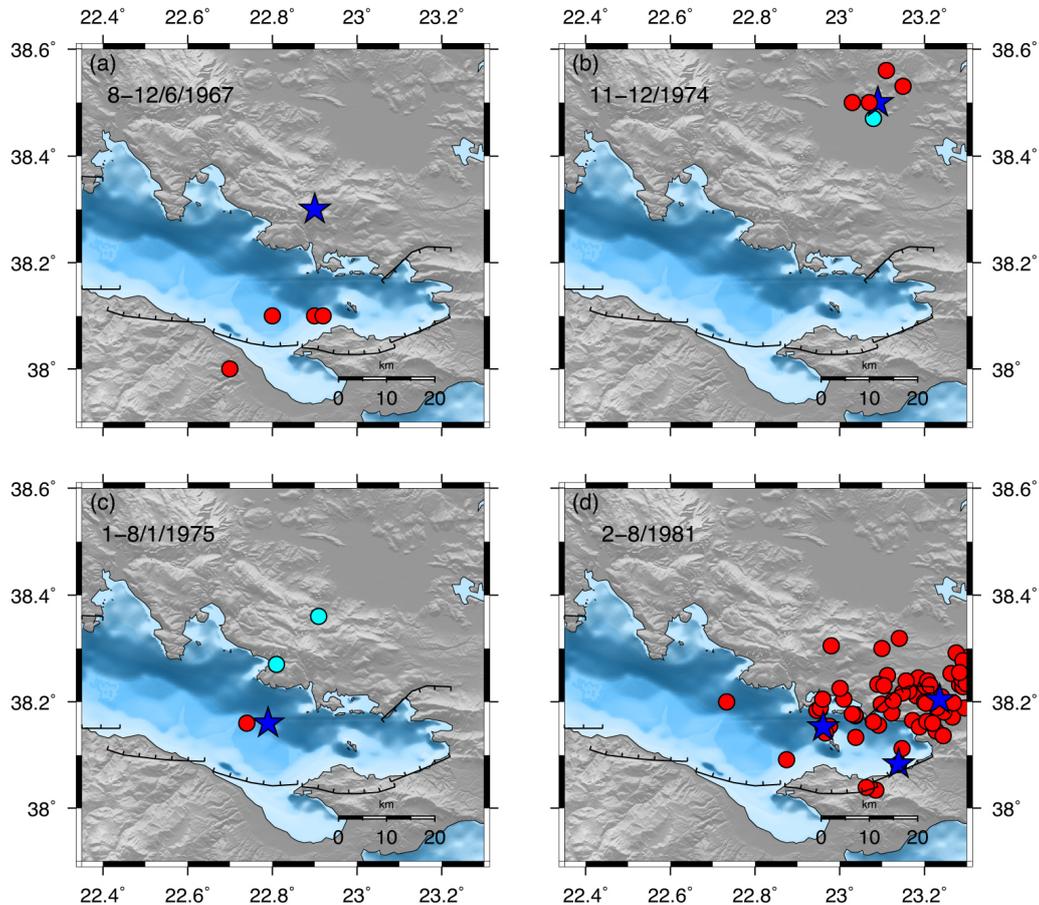
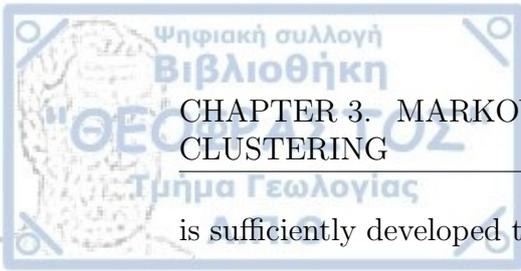


Figure 3.8: Epicentral distribution of the earthquakes that occurred during the sojourn of the process in states 3 and 4 in the eastern subarea during 1964-2017. Stars (blue) represent main shocks and circles correspond to earthquakes that occurred before (cyan) and after (red) the main shock, respectively. Earthquakes associated with the (a) June 1967 seismic sequence, (b) 14 November 1974, (c) 8 January 1975 events and (d) February-March, 1981 seismic sequence.

dow is one more parameter that should be defined with caution. For a robust evaluation of the clustering method, we need a sufficiently long time interval to include a significant number of seismic sequences. The third parameter is the completeness magnitude, which is defined as the magnitude threshold above which we do not miss any earthquake mainly due to deficiencies of the seismic network. It is related to the choice of the time window in the sense that we look for determining a starting date of our catalog when the seismic network



## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

---

is sufficiently developed to achieve a low magnitude of completeness including in this way seismic excitations of a lower scale.

Three major seismic zones in the region of Greece (Figure 3.10), which constitute distinctive seismotectonic units, are chosen to investigate the clustering properties of seismicity with the use of the proposed MAP-DBSCAN method in Section 3.3.2.3. The selection of the three study areas is based on criteria related to the homogeneity of the type of faulting, the comparatively intense continuous seismicity and the existence of seismic excitations during the study period.

### 3.3.2.1 Data

#### Simulated ETAS catalog

For the generation of the simulated ETAS catalog firstly we define the broader spatial and temporal area to avoid the boundary issue, i.e., triggering effect of events outside the target region and from an earlier period, and then the target area  $\Sigma \times [t_{in}, t_f]$  with  $\Sigma \subseteq \Sigma_0$  and  $t_{in} > t_0$  which will be the data set for our study. The broader region is confined by the Greek territory (Figure 1.2) with duration 20 years  $[t_0, t_f] = [0, 20]$  and the target region will be the Corinth Gulf area (Figure 1.3), lasting 18 years  $[t_{in}, t_f] = [2, 20]$ .

We generate the mother events of the catalog using the coordinates from the background seismicity of the Greek earthquake catalog. In this way, they will be distributed in space according to the faulting geometry of the area. We consider earthquakes with  $m_c = 2.5$  during the time period 2011-2019 and implement the Nearest-Neighbor declustering algorithm to separate the clustered from the background events. Then, we produce  $N_{main}$  mother events according to a Poisson distribution with mean value equal to the number of the identified background events. Their coordinates are sampled with replacement from the declustered catalog and a small random factor is added. The occurrence times of the mother events are simulated from a uniform distribution  $U(t_0, t_f)$ ,



## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

where  $t_0 = 0$  and  $t_f = 20$  years and the magnitudes are generated according to the Gutenberg-Richter law (Equation A.7) truncated from the left at the completeness magnitude,  $m_c = 2.5$ , and from the right at the maximum observed magnitude of the instrumental earthquake records in Greece plus a small factor, such that  $m_{max} = 7.8$ . We set the  $b$ -value of the GR law equal to one,  $b = 1.0$ , where  $\beta = b \cdot \log(10)$ .

After the generation of the background events, we simulate their aftershock number, their occurrence times and the locations according to the productivity law (Equation A.2), the Omori law (A.3) and a spatial power law (Equation A.4), respectively, using a typical thinning method. We stop when no more events are triggered inside the target region. In Table 3.8, we give the parameter set that produced the ETAS catalog.

Table 3.8: Parameters used for the generation of the simulated ETAS earthquake catalog with  $m_c = 2.5$ . The target area is  $\Sigma \times [t_{in}, t_f] = [21.3^\circ E - 23.2^\circ E] \times [37.9^\circ N - 38.6^\circ N] \times [2, 20]$ . The number of clustered events is  $N = 4253$  and the number of mother events is  $N_{bg} = 1595$ .

Parameter		Parameter	
$K$	0.1	$d$	$2.41 \times 10^{-5}$
$a$	2.19	$q$	1.805
$p$	1.13	$\gamma$	0.59
$c$	0.024 (days)	$\mu(\text{events/day})$	4.50

### Earthquake catalogs

The earthquake catalogs of CG, CII and NAS, which we denote henceforth as D1, D2 and D3, include 25,595, 24,085 and 21,139 events, respectively, occurring between 2012 and 2019 (Table 3.9). An upgrade on the magnitude determination software of the Greek seismic network happened in early 2011, that led to the significant decrease of the completeness magnitude (Mignan and Chouliaras, 2014). So, we chose 2012 as starting year to avoid significant temporal fluctuations in the magnitude cutoff and include as many earthquakes as possible in our study.

CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Table 3.9: The completeness magnitude,  $M_c$ , for the datasets of the three areas CG, CII and NAS, along with the productivity,  $a$ , and the  $b$ -value of the GR law.  $N$  and  $N_c$  denote the initial number of events and the ones with  $M \geq M_c$ , respectively.

Region	Notation	$N$	$M_c$	$N_c$	$a$	$b$
CG	D1	25,595	1.5	13,043	5.57	0.97
CII	D2	24,085	2.2	6981	5.80	0.88
NAS	D3	21,139	2.1	8328	5.79	0.89

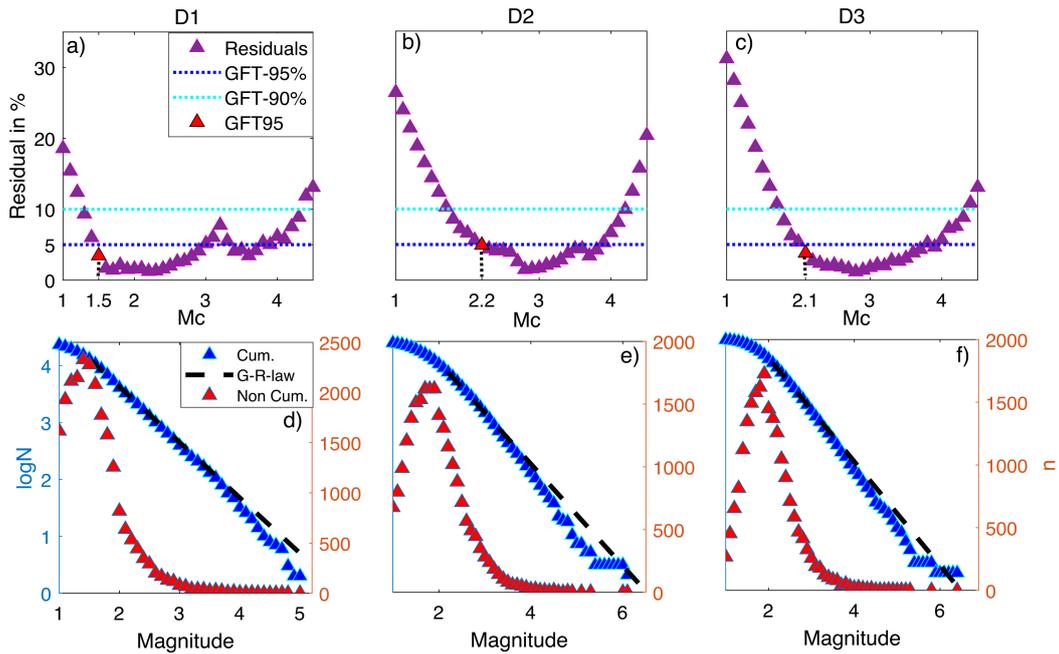


Figure 3.9: (a-c) Residuals (purple triangles) as a function of minimum cutoff magnitude,  $M_c$ , for the D1, D2 and D3 data sets, respectively. Blue and cyan dotted horizontal lines indicate the 10% and 5% residual thresholds, respectively.  $M_c$  (red triangle) is found as the first magnitude cutoff at which the confidence 95% is reached. (d-f) incremental (red triangles) and cumulative frequency (blue triangles) as a function of magnitude. The dashed black line is the GR law fit according to the GFT method with  $M_c = 1.5, 2.2, 2.1$  for data sets D1, D2 and D3, respectively.

For the determination of the completeness magnitude, we applied the Goodness-of-Fit (GFT) method (Wiemer and Wyss, 2000). It is based on the assumption that a power law, known as Gutenberg-Richter (GR) law (Gutenberg and Richter, 1944),  $\log N = a - b \cdot M$ , where  $N$  the cumulative number of events with magnitude larger than or equal to  $M$ , can approximate the frequency

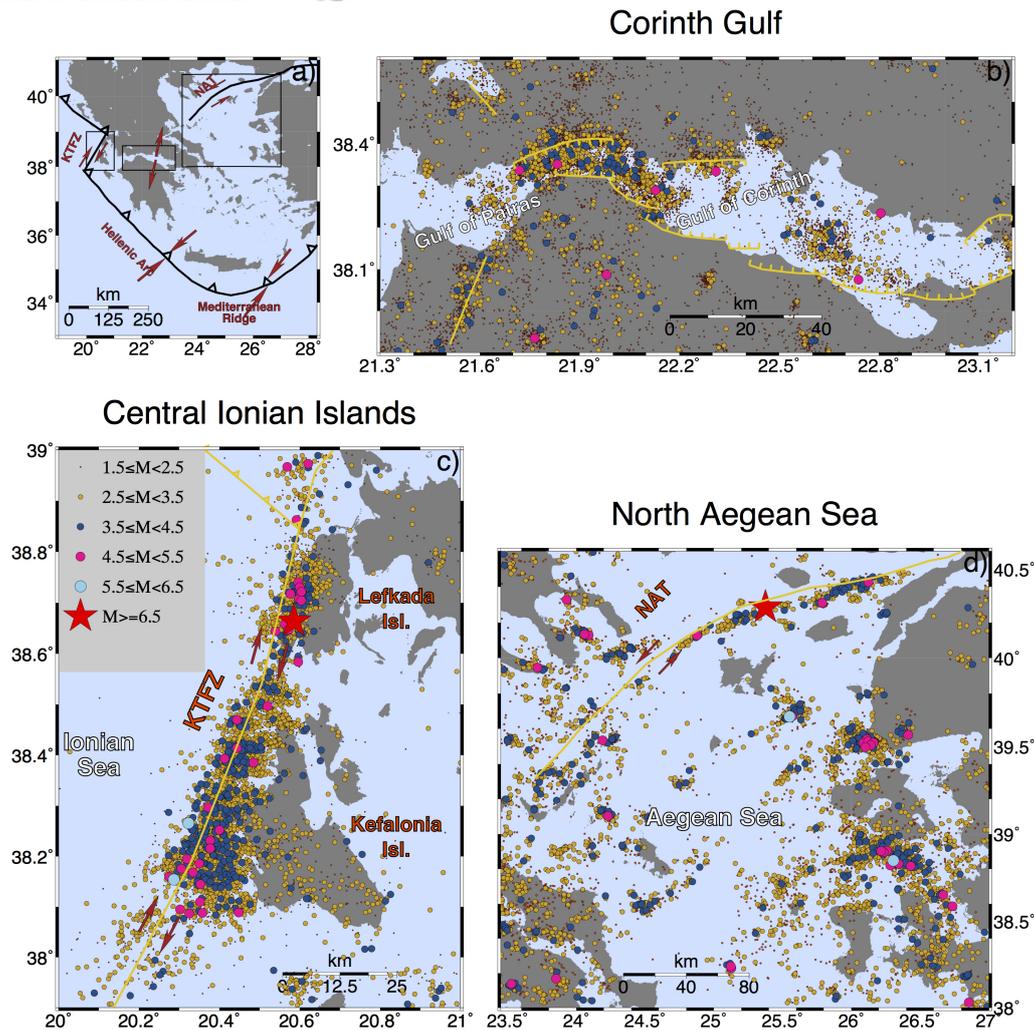


Figure 3.10: Maps of the study areas depicting seismicity along with major faults (yellow lines). (a) The area of Greece with the main structures (thick black lines), such as the Hellenic Arc (thick black line with triangles), the North Aegean Trough, and the Kefalonia dextral Transform Fault Zone. Red arrows represent the kinematics along major fault zones as well as the main deformational pattern in the Corinth Gulf, characterized by extension. (b) The area of Corinth Gulf where the major faults are shown (yellow lines) along with seismicity during 2012–2019. (c) The area of Central Ionian Islands where the Kefalonia Transform Fault Zone and the collision front are shown (yellow lines) along with seismicity during 2012–2019. (d) The area of North Aegean Sea where the NAT is traced (yellow line) along with seismicity during 2012–2019. The legend is common for the three study areas.

magnitude distribution of earthquakes. In particular, the differences, which are called residuals, between the observed and the synthetic frequency-magnitude

distribution are computed for increasing magnitude bins as threshold values. The completeness magnitude corresponds to the first magnitude bin at which the difference falls under a predefined error threshold, usually 5%. The parameters of the synthetic distribution,  $a$  and  $b$ , are calculated by means of the maximum likelihood method (Aki, 1965) and the residuals,  $R_i$ , for each magnitude bin,  $M_i$ , between the observed and synthetic distributions are given by  $R_i = 100 - (\frac{\sum_{M_i}^{M_{max}} |B_i - S_i|}{\sum_i B_i} \cdot 100)$ , where  $B_i$  and  $S_i$  are the observed and expected cumulative number of events with  $M \geq M_i$ , respectively. Figures 3.9a–c show the residuals for the three data sets and Figures 3.9d–f present the cumulative and noncumulative frequency and the fitted GR law for the corresponding complete data sets. The magnitude thresholds and the corresponding estimated  $b$ -values for the three data sets are equal to  $M_c = 1.5, 2.2, 2.1$ , and  $b = 0.97, 0.88, 0.89$  with 13, 043, 6981, 8328 events (Table 3.9), respectively. The epicenters are shown in the map of each study area, for Corinth Gulf in Figure 3.10b, for Central Ionian Islands in Figure 3.10c and for North Aegean Sea in Figure 3.10d.

### 3.3.2.2 Evaluation on synthetic catalog

We implemented the GK, RB, NN and MAP-DBSCAN clustering algorithms to the simulated ETAS catalog that lasts 18 years and is located within the Corinth Gulf area. For the window-based method we used three different temporal and spatial intervals, given by Equations (B.1) (GK1), (B.2) (GK2) and (B.3) (GK3), respectively. They resulted to 341, 223, 319 earthquake clusters and 596, 152, 1821 single events, respectively.

For the Reasenberg algorithm, we used the ZMAP tool (Wiemer, 2001) where we provided as input the simulated ETAS catalog and we adopted 3 different sets of parameters, notated as RB1, RB2 and RB3, respectively, which are given in Table 3.10. RB1 (Table 3.10) corresponds to the original parameters proposed in Reasenberg (1985). In the second set, RB2, we extend the spatial zone by increasing the factor  $r_{fact}$  from 10 to 20 km, whereas, in the

CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

third set, RB3, we also extend the temporal window modifying the parameters  $\tau_{min}$  and  $\tau_{max}$  (Table 3.10). As a result we have 189, 188, 139 earthquake clusters and 2437, 2324, 2428 single events, respectively.

Table 3.10: Input parameters for the Reasenberg clustering algorithm. The first row corresponds to the standard parameter set (Reasenberg, 1985).

PS	$\tau_{min}$	$\tau_{max}$	$p_1$	$x_k$	$x_{meff}$	$r_{fact}$
RB1	1	10	0.95	0.5	2.5	10
RB2	1	10	0.95	0.5	2.5	20
RB3	0.5	20	0.95	0.5	2.5	20

The Nearest-Neighbor algorithm requires as input only two free parameters, the spatial fractal dimension,  $d_f$ , and the  $b$  value (Equation B.4), which are considered equal to  $d_f = 1.51$  and  $b = 1.0$ , respectively. The logarithm of the separation distance is equal to  $\log \eta_0 = -5.04$ , based on the intersection of the two modes in the 1D density distribution of the logarithmic spatiotemporal distances,  $\eta_{ij}$ , computed through Equation B.5 (Figure 3.11). The NN algorithm detected 311 earthquake clusters and 1648 single events.

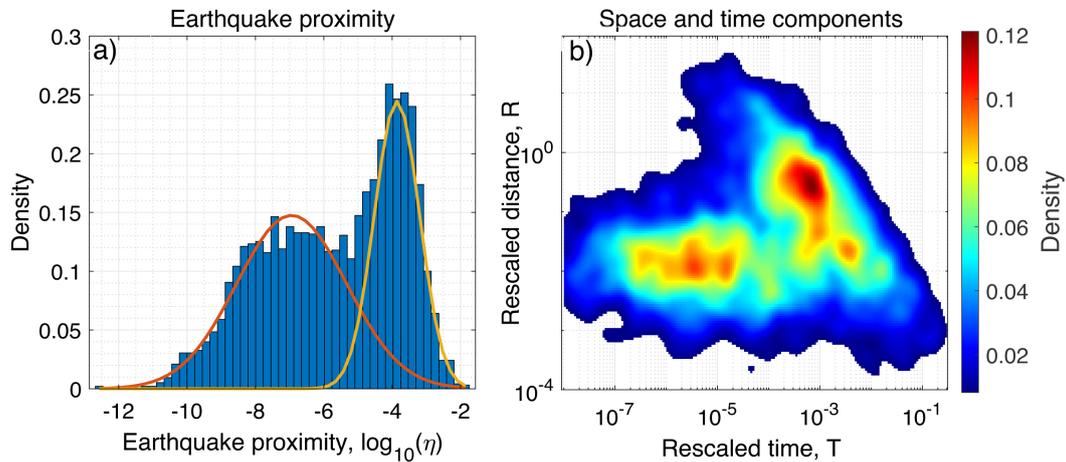


Figure 3.11: Distribution of the NN distances among all pairs of earthquakes of the ETAS synthetic catalog. (a) 1D density distribution of  $\log \eta$ , with estimated Gaussian densities for clustered (yellow) and background (orange) components. (b) 2D joint distribution of rescaled space and time distances.

Next, we implemented the MAP-DBSCAN procedure. In the first step we need to choose the optimal number of occurrence rates for the MAP model.

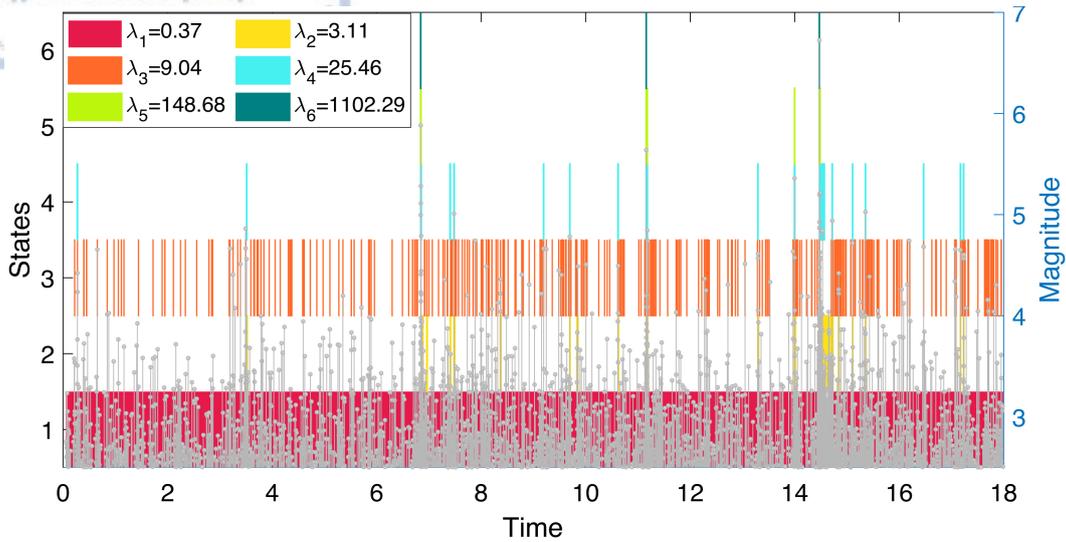
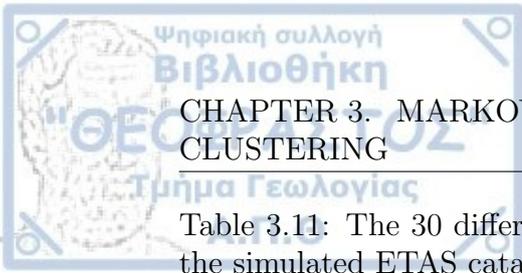


Figure 3.12: Most probable path of the hidden states of the model along with the magnitude of events as a function of time (grey vertical lines). Each color is assigned to a different state  $i$  with seismicity rate  $\lambda_i$ . The rate threshold,  $\lambda_{thr}$ , is set equal to  $\lambda_1 = 0.37$ .

We fitted MAP models up to 7 states and the BIC criterion (Equation 2.17) suggested a MAP with 6 states. The most probable sequence of the hidden states of the MAP model is computed through the state probabilities given in Equation 2.14 and is shown in Figure 3.12. From Figure 3.12 is evident that state 1 with  $\lambda_1 = 0.37$  corresponds to the background rate of the data set, so it was set as the rate threshold  $\lambda_{thr} = \lambda_1$  for the initial separation of seismicity into potential clusters.

Next, different temporal windows,  $(T, dt)$ , are tested for merging the potential clusters. The most rigorous selection is with  $T = 0$  and  $dt = 0$  and the most loose one is with  $T = 14$  days (potential clusters within 14 days are merged into one) and  $dt = 7$  (events  $\pm 7$  days from the potential cluster are included), respectively. Finally, the DBSCAN algorithm is implemented for 5 different distance thresholds ( $\epsilon = \{2.5, 7.5, 10, 12.5\}$  km). The minimum number of events is set to  $N_{pts} = 2$  for a better comparison with the other methods where clusters with at least 2 events can be defined. In Table 3.11 we present details on the parameter tuning.



## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Table 3.11: The 30 different parameter sets used for the clusters detection of the simulated ETAS catalog.

$\epsilon$	$N_{pts}$	PS	$T$	$dt$	PS	$T$	$dt$
		1-5	0	0	16-20	0	7
[2.5 5 7.5 10 12.5]	2	6-10	7	0	21-25	7	7
		11-15	14	0	26-30	14	7

The method seems rather insensitive to the parameter selection. In particular, Figure 3.13 presents the Jaccard index values given by Equations 3.2 and 3.3, that describe the efficiency of the method to correctly reconstruct the initial clusters ( $J_1$ ) as well as to identify the single events ( $J_2$ ), respectively. We observe that the Jaccard index values are quite stable with small fluctuations, apart from the smallest upper-distance cutoff,  $\epsilon = 2.5$  km, which seems inadequate to capture the spatial correlations among the events. Furthermore, the contribution of the temporal constraints to the clustering procedure seems negligible, with the exception of the two peaks for PS12 and PS27. This is an indicator that the MAP model has already achieved a sufficient separation between background and triggered seismicity with the six embedded rates,  $\lambda_i$ ,  $i = 1, \dots, 6$ , of the model.

In Table 3.12, we show the Jaccard index values ( $J_i, i = 1, 2$ ) (Equations 3.2, 3.3) for all the clustering algorithms. In particular, for the MAP-DBSCAN algorithm we show the one with the best results in terms of the Jaccard index (MAP-DBSCAN27).

Table 3.12: The  $J_i, i = 1, 2$ , values for 3 parameter sets (PS) of the Reasenber and Garden-Knopoff algorithms, respectively, and for the corresponding MAP-DBSCAN and Nearest-Neighbor methods.

PS	RB1	RB2	RB3	GK1	GK2	GK3	MAP-DBSCAN (PS27)	NN
$J_1$	0.530	0.593	0.648	0.382	0.397	0.585	0.627	0.756
$J_2$	0.612	0.630	0.617	0.418	0.192	0.676	0.647	0.727

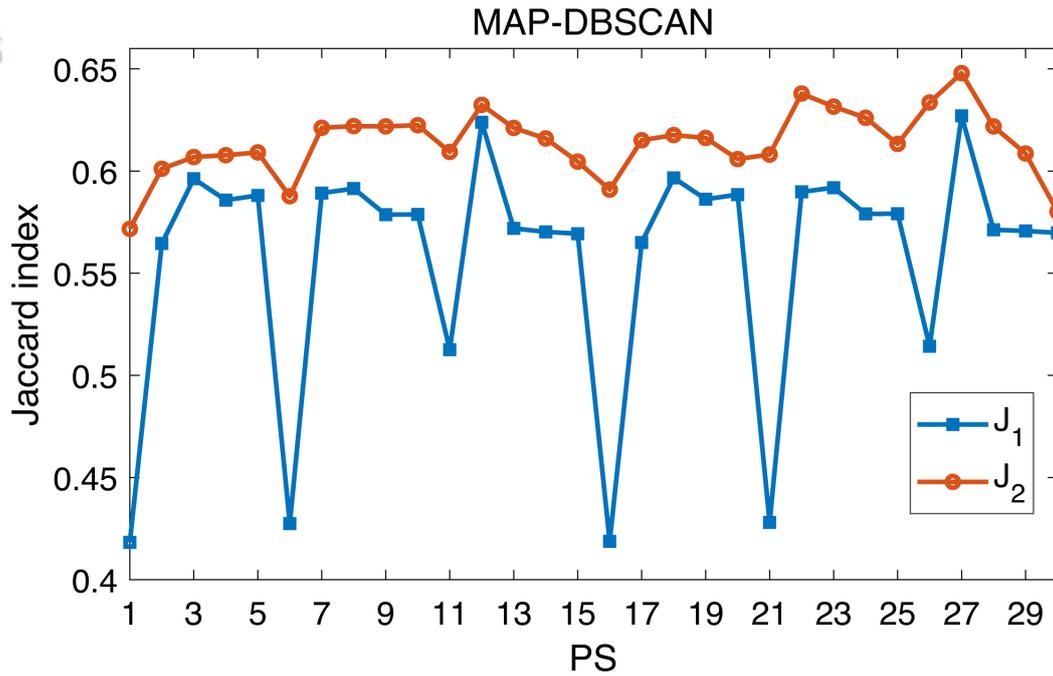


Figure 3.13: The Jaccard index values,  $J_1$  with blue and  $J_2$  with orange color, respectively, for all the input parameters of the MAP-DBSCAN method.

The Nearest-Neighbor method shows the best performance among the testing algorithms in the reconstruction of the clusters ( $J_1 = 0.756$ ) and in the detection of the mother events ( $J_2 = 0.727$ ). This is also evident by its cumulative number of background seismicity (purple line in Figure 3.14a), which is the closest one to the initial catalog (dotted black line). The temporal evolution of background seismicity is shown in Figures 3.14b–f across the longitude for ease of reading as west–east normal faults dominate the Corinth Gulf area. For the NN method the space-time evolution of the declustered seismicity is smooth without large gaps, although there is a significant concentration of events between the 7th and 8th year of the catalog, which is also persistent in both RB2 and GK3 methods (orange ellipses in Figures 3.14d–f) and less apparent on MAP-DBSCAN method (orange ellipse in Figure 3.14c). The high efficiency of the NN method is probably related to the metric it uses, which is similar to the ETAS one with  $\lambda_j(t_i, x_i) = (t_i - t_j)^{-1} r_{ij}^{-d_j} 10^{bm_j}$  and  $c = 0$ ,  $p = 1$ ,  $d = 0$ ,  $q = d_f$  and  $a = b$ . The windowing technique seems to overestimate the temporal and

spatial windows, since it removes large amounts of seismicity (blue ellipse in Figure 3.14f), in accordance with previous results (Peresan and Gentili, 2020). The same gap between the 14th and 15th year of the catalog is also evident in the background seismicity from the MAP-DBSCAN method, however, it is smaller and some sparse seismicity is left (blue ellipse in Figure 3.14c). On the other hand, Reasenberg's declustered catalog has more events than any other method (pink, magenta and green line, Figure 3.14a) and significant concentrations of events are visible in the space-time evolution of the background seismicity (orange and purple ellipses in Figure 3.14e).

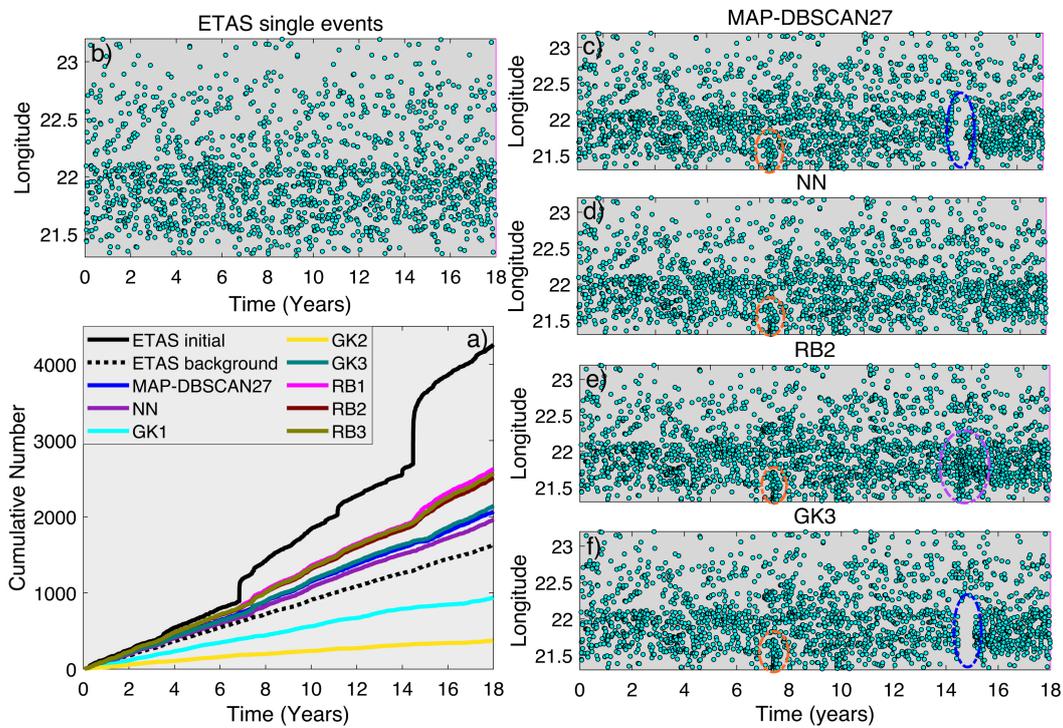


Figure 3.14: (a) Cumulative number of background events for each algorithm, the initial ETAS catalog with black color and the mother events of the ETAS synthetic catalog with the black dotted line. The space-time evolution (b) of the initial ETAS catalog and of the background seismicity for the four best algorithms, (c) MAP-DBSCAN27 ( $J_2 = 0.647$ ), (d) NN ( $J_2 = 0.727$ ), (e) RB2 ( $J_2 = 0.630$ ), (f) GK3 ( $J_2 = 0.676$ ). Colored ellipses stand for large gaps and significant concentration of events.

Best overlapping among the true,  $\mathbf{X}$ , and the estimated partition,  $\mathbf{Y}$ , does not mean necessarily the best detection for the declustered seismicity. For

instance, the GK3 partition is characterized by a lower index,  $J_1 = 0.585$ , than the MAP-DBSCAN27 partition,  $J_1 = 0.627$ , however, its declustering catalog is more accurate ( $GK3-J_2 = 0.676 > J_2 = 0.647$ -MAP-DBSCAN27). Nevertheless, both indexes combined ( $J_1 + J_2$ ), the MAP-DBSCAN partition shows a higher efficiency from the rest of the algorithms, except from NN.

### 3.3.2.3 Potential clusters identification in three seismic zones of Greece

We applied the MAP-DBSCAN procedure to the data sets of the three major seismic zones, Corinth Gulf, Central Ionian Islands and North Aegean Sea for earthquakes with  $M \geq 1.5$ ,  $M \geq 2.2$  and  $M \geq 2.1$ , respectively, that occurred during 2012-2019. We fit MAP models from two up to seven states for each earthquake sub-catalog, computationally a very demanding process as the number of states increases, especially for large data sets such as D3 with 13043 events. The BIC is used as a metric to determine the optimal number of hidden states for each model as it tends to the "true" model for large data sets (Burnham and Anderson, 2002). The lowest BIC values correspond to six, seven and again six states of the MAP model for the D1, D2 and D3 data sets, respectively.

Next, we evaluate the most probable path of the hidden states of the models though the state probabilities  $p_i(t) = p(J_t = i)$  for  $i = 1, \dots, K$ , given in Equation 2.14. Figures 3.15a–c illustrate the transitions among the states for the data sets D1, D2 and D3, respectively. The colored box at each temporal interval  $t_k \leq t < t_{k+1}$  indicates the state with the maximum probability at the certain time and the legend contains its corresponding occurrence rate.

The temporal patterns of dataset D1 indicate the dominance of state 2 (yellow color, Figure 3.15a) with occurrence rate  $\lambda_2 = 3.01$  events/day for almost the entire period. Nevertheless, there is a slight decrease in the occurrence of earthquakes ( $\lambda_1 = 1.23$  events/day) in the second part of the catalog, starting from 02/2016 with transitions to state 1 (red color, Figure 3.15a) until almost

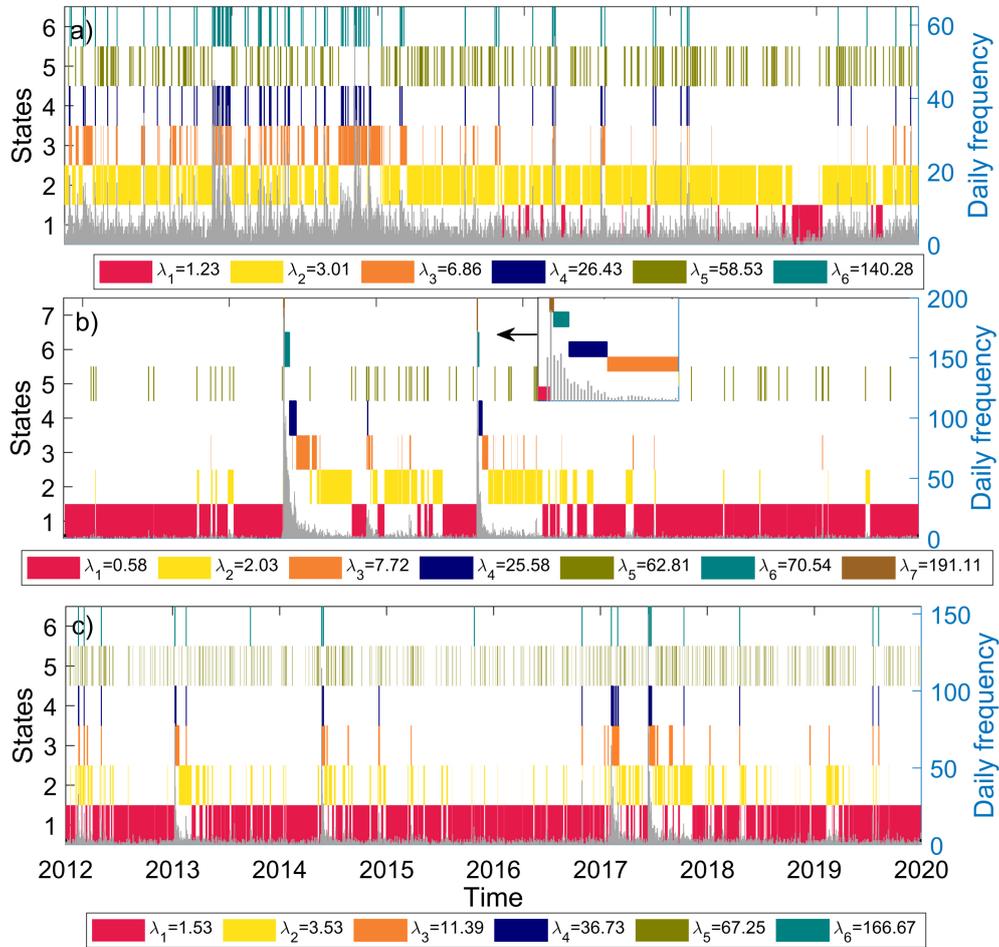


Figure 3.15: Most probable path of the hidden states of the model along with the daily frequency of events (gray vertical bars) with (a)  $M \geq 1.5$  for D1, (b)  $M \geq 2.2$  for D2 and (c)  $M \geq 2.1$  for D3 datasets, respectively. Each color is assigned to a different state  $i$  with seismicity rate  $\lambda_i$ . Inset magnifies the transitions among the states, which are otherwise difficult to visualize due to the short sojourn times compared to the study period. The rate threshold,  $\lambda_{thr}$ , is set equal to  $\lambda_2 = 3.01$ ,  $\lambda_1 = 0.58$ ,  $\lambda_1 = 1.53$  for the D1, D2 and D3 datasets, respectively.

the end of the catalog in 12/2019. This is probably related to the lack of seismic sequences during the last part of the study period compared to the previous intense seismic activity especially during the period 2013–2014 in the western subarea of the CG (Kapetanidis et al., 2021). The rate threshold is set equal to  $\lambda_{thr} = \lambda_2$ , which we consider as the background rate during the study period.

The seismicity of the CII area is dominated by the two major sequences during the study period, the 2014 Kefalonia doublet ( $M_w$ 6.1 and  $M_w$ 6.0) (Karakostas et al., 2015) and the  $M_w$ 6.5 2015 Lefkada earthquake sequence (Papadimitriou et al., 2017). States 7 (brown), 6 (dark cyan), 4 (dark blue), 3 (orange) and 2 (yellow) in Figure 3.15b are clearly associated with the aftershock evolution of the two sequences, essentially, they approximate the Omori temporal distribution. Background seismicity is described by state 1 (red) with occurrence rate  $\lambda_1 = 0.58$  events/day, which we set as rate threshold for the primary classification of seismicity into potential clusters.

Finally, dataset D3 also contains some major sequences, the 2013  $M_w$ 5.8 North Aegean (Karakostas et al., 2014), the 2014  $M_w$ 6.9 Samothraki (Saltogianni et al., 2015) and the 2017  $M_w$ 6.4 Lesvos earthquake sequences (Papadimitriou et al., 2018), whose aftershock temporal distribution is approximated by states 6 (dark cyan), 4 (dark blue), 3 (orange) and 2 (yellow) of the model (Figure 3.15c). The rate threshold value is set equal to  $\lambda_{thr} = \lambda_1$ .

In general, we observe significant variations in the temporal evolution of the seismic excitations between the CG and the CII and NAS areas. Figure 3.15 illustrates that the daily frequency of events (grey vertical bars) during seismic sequences in CII and NAS is decreasing in time, typical of main shock–aftershock sequences, whereas in CG we observe large fluctuations in the daily frequency, common for earthquake swarms, as in 2014 when multiple seismic excitations occurred in the western subarea of CG.

Consecutive events above the rate threshold,  $\lambda_{thr}$ , are classified into groups which are called potential clusters. The three data sets, D1, D2 and D3 consist of 314, 49 and 281 potential clusters, respectively.

### 3.3.2.4 Sensitivity analysis

We now provide more details on the choice of the parameters and how they affect the spatiotemporal evolution of background seismicity. We implemented the clustering procedure, MAP-DBSCAN, for 16 different combinations of pa-

## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

parameters which are shown in Table 3.13. In particular, we tested four different sets of temporal constraints,  $(T, dt)$ , for the merging of the potential clusters of the three data sets and then we applied the DBSCAN algorithm to the merged clusters in order to separate them based on their spatial density. The minimum number of neighbors for the determination of a cluster is set equal to 4 ( $N_{pts} = 4$ ) for avoiding cases with few events. This is an appropriate choice for two-dimensional data according to Ester et al. (1996).

Table 3.13: The 16 tested parameter of MAP-DBSCAN method for the three datasets D1, D2 and D3.

$\epsilon$	$N_{pts}$	$PS$	$T$	$dt$	$PS$	$T$	$dt$
[2.5 5 7.5 10]	4	1	0	0	3	0	5
		2	5	0	4	5	5

For the determination of the distance threshold  $\epsilon$ , we computed the  $k$ -distances between events assigned to the same potential cluster, since the DBSCAN algorithm is implemented in events that have been already grouped into clusters based on their temporal proximity. This is a procedure proposed by Ester et al. (1996), which is commonly used to constrain the distance threshold (Petersen et al., 2021). In particular, for each event included in the potential cluster, its  $k$ -nearest neighbor is computed and plotted in ascending order. If we choose an arbitrary event,  $i$ , set the distance threshold  $\epsilon$  to  $k$ -dist( $i$ ) and the parameter  $N_{pts}$  to  $k$ , all events with an equal or smaller  $k$ -dist value will become core points, in other words, they will be assigned into a cluster. Ester et al. (1996) proposed as best  $\epsilon$  value the one that corresponds to a change in the slope of the curve, as corner points indicate a change in the degree of correlation among events. For  $k = 4$ , which corresponds to the minimum number of neighbors ( $N_{pts}$ ), gradient changes in the slope range between 2.5 and 10 km in the data sets of both CG (Figure 3.16a) and NAS (Figure 3.16c) areas, whereas for the CII area (Figure 3.16b), changes in the slope of the curves initiate slightly sooner (below 2.5). The minimum one is chosen as

equal to  $\epsilon = 2.5$  in order to also ensure that the location errors of the catalog are considerably lower.

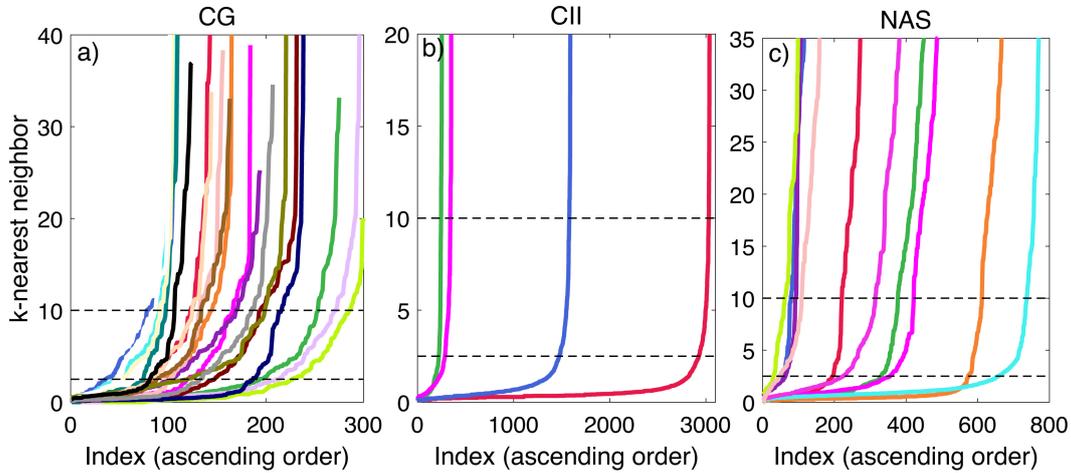


Figure 3.16: The k-nearest neighbor plot of the potential clusters with  $N \geq 100$  events in (a) CG (b) CII and (c) NAS. Black horizontal dashed lines indicate the range of  $\epsilon$  values given as input to the DBSCAN algorithm and each color corresponds to a potential cluster.

For the 16 different realizations of the clustering algorithm, MAP-DBSCAN, we investigated the spatiotemporal properties of the background seismicity, i.e., events that have not been assigned to a cluster. Figure 3.17 presents the cumulative number of the declustered seismicity for each set of parameters along with the initial data sets. Peaks and pronounced concavities in the cumulative curves are indicators of triggered seismicity wrongly assigned as background and vice versa. In data sets D1 and D2 we observe such concaves for thresholds  $\epsilon \geq 5$  km and a rather stable curve for  $\epsilon = 2.5$  km (Figure 3.17a–h), suggesting that events are correctly separated as background and triggered ones. Therefore, the distance threshold is set to  $\epsilon = 2.5$  km, for both data sets. In data set D3, Figure 3.17i–l show that the curves with  $\epsilon \geq 7.5$  km exhibit large concaves, indicating that background seismicity is incorrectly assigned to clusters. For the smallest threshold  $\epsilon = 2.5$  km, some small peaks appear and thus the  $\epsilon = 5$  km as the optimal value was selected. Data set D3 contains offshore seismicity in the NAS area, with probably higher location

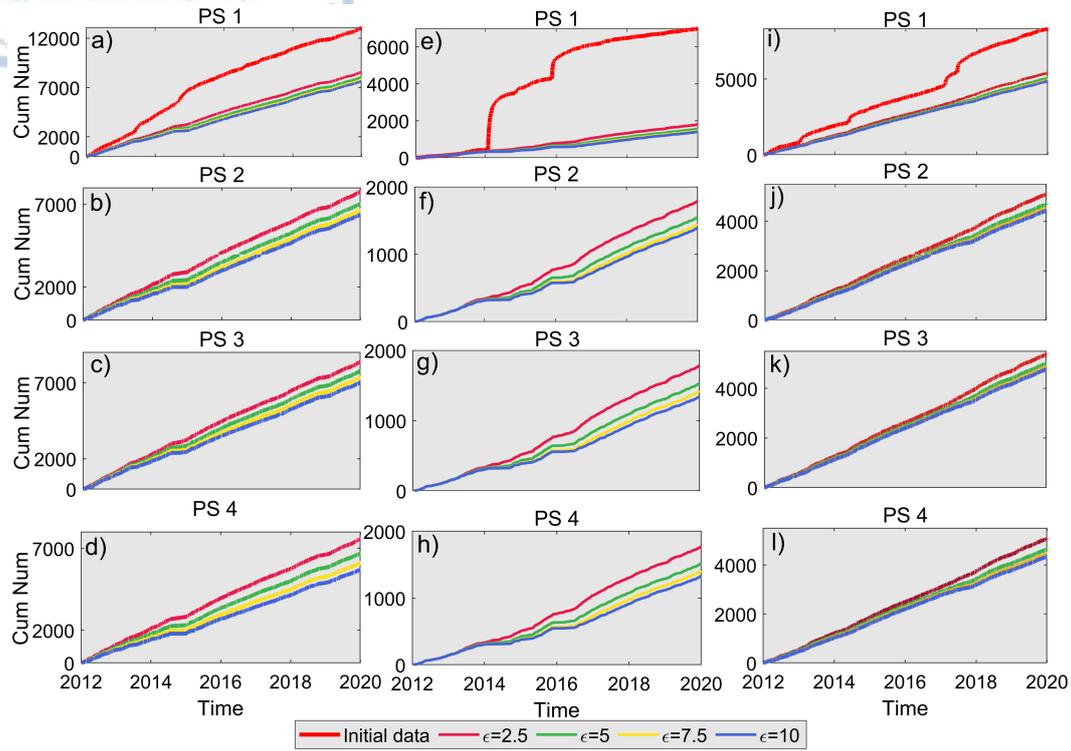


Figure 3.17: Cumulative number of the initial data sets (red line) and cumulative number of background seismicity for each parameter set (PS1-PS4) and for four different distance thresholds ( $\epsilon = 2.5, 5, 7.5, 10$  km). (a–d) Data set D1, (e–h) dataset D2 and (i–l) dataset D3.

errors. This supports our choice for a larger distance threshold.

To further explore the differences between the spatiotemporal evolution of the declustered catalogs, the space-time pattern of the background events is examined, comparing the full and the declustered catalogs. In data set D1, a persistent gap of seismicity appears during the second half of 2014, independently of the chosen temporal constraints, associated with the two large earthquake swarms in that period (Duverger et al., 2018). Due to the intense seismic activity during 2013–2014 in the western Corinth Gulf (Kapetanidis et al., 2021; Michas et al., 2021), the classification of seismicity into clusters becomes more complicated, so we have chosen a rather conservative parameter set,  $PS3$ , with  $T = 0$ . In this way, we avoid merging distinct clusters that are spatio-temporally close to each other. Figure 3.18a shows the space-time evolution of the declustered catalog that corresponds to the final parameter

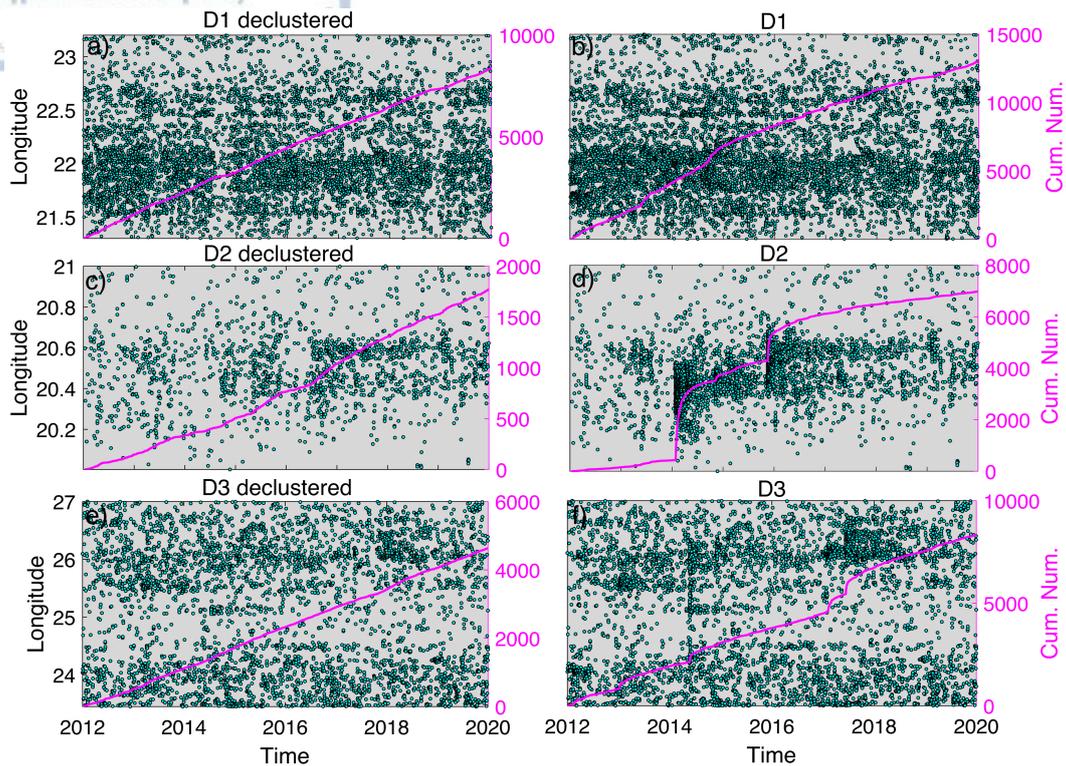


Figure 3.18: Space-time evolution of the background and initial seismicity for data set (a,b) D1, (c,d) D2 and (e,f) D3. Purple lines denote the cumulative number of events.

set. The main seismic excitations present in Figure 3.18b are detected, while preserving the patterns of the background seismicity. In dataset D2, the results are quite similar for all the tested temporal constraints, and for this reason, we adopted parameter set  $PS4$  with  $T = 5$  days, which is a more loose constrain. It is more likely for seismic excitations close in time to be part of the same main shock–aftershock sequence, due to the two major sequences that dominate in the study period. In the initial dataset (Figure 3.18c), the two major sequences are visible, whereas they are removed after the implementation of the clustering algorithm, while preserving the main patterns of background seismicity (Figure 3.18d). Finally, for the NAS area, the differences over the temporal constraints seem negligible, therefore, we chose parameter set  $PS4$ . Figure 3.18e illustrates a standard scattering of the background seismicity in space without gaps and high-density areas, whereas the main seismic sequences visible in Figure 3.18f

have been identified.

### 3.3.2.5 Cluster analysis in the three seismic zones

Table 3.14 gives the final parameter set of the clustering procedure for each data set based on the analysis in Section 3.3.2.4 and a summary on the statistics of the detected clusters. In the CG area, we identified the largest number of seismic clusters (255) due to the increased detectability of micro-seismicity (low completeness magnitude threshold), however, they are short in size ( $\bar{n} = 18.28$ ) and duration ( $\bar{\tau} = 12.50$ ). Conversely, the CII area is characterized by a small number of seismic clusters (45) but with large mean size ( $\bar{n} = 118.43$ ) and duration ( $\bar{\tau} = 54.60$ ). The clustered seismicity is prevalent (75%), whereas in CG and NAS, the background component is more dominant than clustered seismicity with 64% and 56%, respectively (Table 3.14). In CG, this is explained by the lack of large main shocks during the study period and the occurrence of few moderate events, the largest number with  $M = 5.2$ .

Table 3.14: Cluster statistics and the parameter set of the clustering algorithm for the three data sets.  $N_{clust}$  corresponds to the number of clustered events and  $N_{bg}$  to the background seismicity frequency.  $\bar{\tau}$  and  $\bar{n}$  are the mean duration in days and size of the clusters, respectively.

Dataset	( $T, dt, \epsilon, N_{pts}$ )	$N_{clust}$	$N_{bg}$	# clusters	$\bar{\tau}$	$\bar{n}$
D1	(0, 5, 2.5, 4)	4662 (36%)	8381 (64%)	255	12.50	18.28
D2	(5, 5, 2.5, 4)	5221 (75%)	1770 (25%)	45	54.60	118.43
D3	(5, 5, 5, 4)	3688 (44%)	4640 (56%)	187	15.08	19.72

### Corinth Gulf Area

The majority of the clusters are located on the western subarea where 22 out of 27 clusters with  $N \geq 30$  occurred. The main activity is located offshore Aigion and also offshore Psathopyrgos fault (Figure 3.19a). The activity of the eastern subarea consists of smaller clusters that are mainly concentrated offshore Xylokastro and Perachora faults, as well as near Itea Gulf (Figure 3.20a).

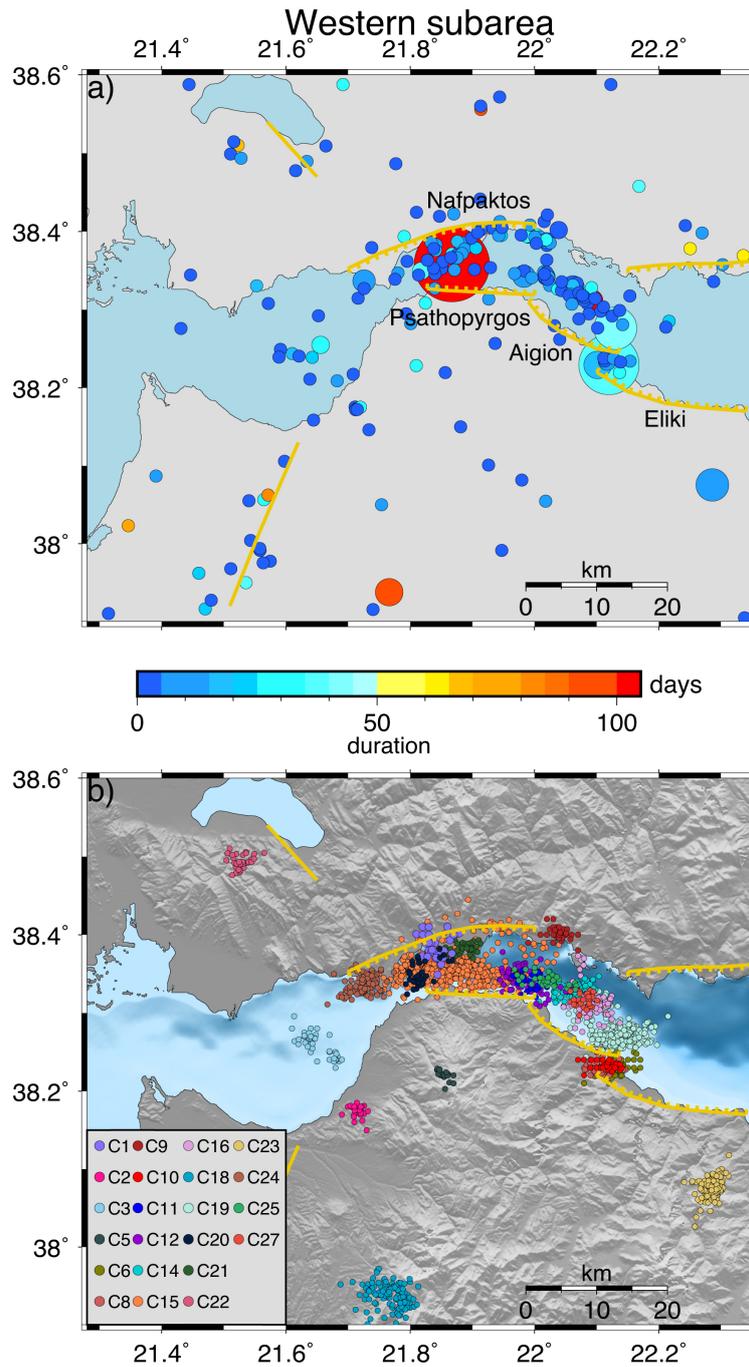


Figure 3.19: (a) Spatial distribution of the centroids of the identified clusters for the western subarea of Corinth Gulf along with major faults (yellow lines). The size of the circles is proportional to the earthquake number in each cluster, whereas the duration is represented by the color scale. (b) Spatial distribution of the clusters with  $N \geq 30$  events. The index of each cluster is provided in the inset box.

The western subarea of the Corinth Gulf is characterized by rich seismic activity, especially in 2013–2014, when 13 out of the 22 clusters with  $N \geq 30$  occurred. One of the major detected sequences is the 2013 Aigion swarm ( $C6$  in Figure 3.19b) which initiated on 21 May 2013 with a plethora of small events ( $M \geq 1.5$ ) and several bursts associated with earthquakes of magnitudes ranging between 3.3–3.7 (Figure C.2) (Kapetanidis et al., 2015; Mesimeri et al., 2016). Two distinct excitations followed ( $C8$  and  $C10$  in Figure 3.19b) in accordance with the ones observed by Michas et al. (2021). The first cluster began on 7 July with some activity prior to the  $M = 3.7$  event on 15 July, 2013, and lasted until 27 August, 2013 (Figure C.2). The second half of 2014 is also a well-studied period with intense seismic activity. Five clusters with  $N \geq 30$  are detected ( $C15$ ,  $C16$ ,  $C18$ ,  $C19$  and  $C20$ ) in the western subarea, including the offshore Aigion  $M4.8$  earthquake on 7 November 2014, associated with  $C19$  (Figure 3.19b), and the  $M4.6$  event on 21 September 2014, associated with the earthquake swarm located between Nafpaktos and Psathopyrgos (Kapetanidis, 2017) ( $C15$  in Figure 3.19b). Persistent activity since 22 July 2014 is also observed offshore Aigion ( $C16$ ), close to the earthquake swarm,  $C15$ , which began on 7 November 2014 (Figure C.4). In 2012, fewer clusters are observed, mostly during the first semester, with three clusters comprising  $N \geq 30$  events,  $C1$ ,  $C2$  and  $C3$ , and a plethora of smaller ones (Figures 3.19b and C.1). Between November 2013 and July 2014, the activity is sparse with three relatively large clusters,  $C11$ ,  $C12$  and  $C14$  (Figures 3.19b and C.3). Six more clusters with  $N \geq 30$  are observed until the end of 2017 ( $C21$ ,  $C22$ ,  $C23$ ,  $C24$ ,  $C25$ ,  $C27$ , Figure 3.19b).

The eastern subarea is characterized by more sparse activity. A major seismic sequence, Offshore Perachora ( $C4$  in Figure 3.20b), is detected, including two sub-sequences, the first initiated on 22 September and the second on 30 September 2012 (Figure C.5). Two relatively large clusters,  $C13$  and  $C17$ , are observed near Itea Gulf; the former lasted almost two weeks at the end of March, 2014, and the latter—almost three months between August and

October 2014 (Figures 3.20 and C.6).

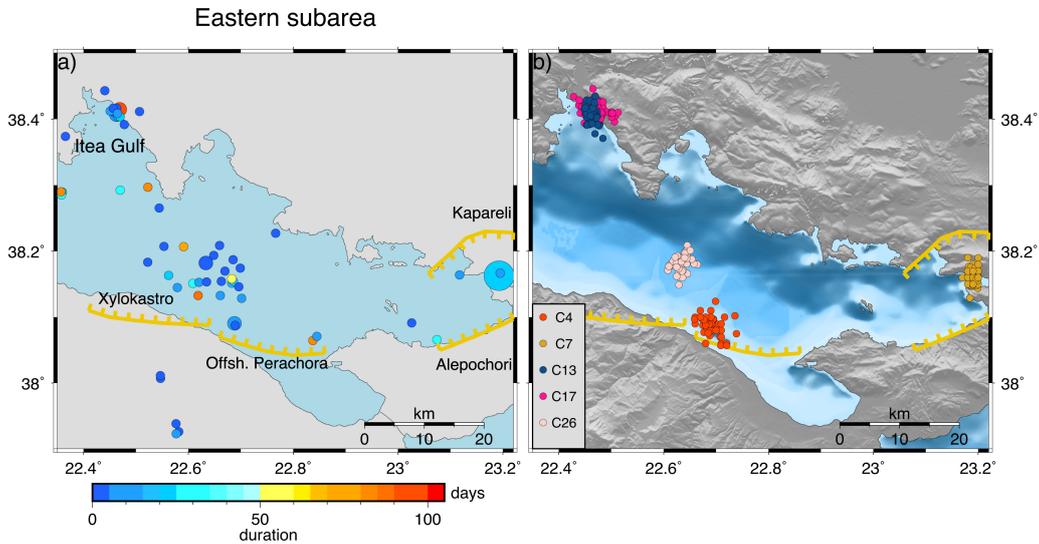


Figure 3.20: (a) Spatial distribution of the centroids of the identified clusters for the eastern subarea of the Corinth Gulf along with the major faults (yellow lines). The size of the circles is proportional to the earthquake number in each cluster whereas the duration is represented by the color scale. (b) Spatial distribution of the clusters with  $N \geq 30$  events. The index of each cluster is provided in the inset box.

### Central Ionian Islands area

The seismicity of CII is dominated by the 2014, Kefalonia and 2015, Lefkada major main shock–aftershock sequences, each one comprising 2829 and 1396 events, respectively. Essentially, 4225 out of the 5221 clustered events belong to these sequences (Table 3.14). Furthermore, 45 clusters are detected in total with the main activity concentrated along the KTFZ (Figure 3.21a).

The two main shocks of sequence *I1* (Figure 3.21b) with  $M = 6.1$  and  $M = 6.0$  occupy the southern and the central part of the onshore area of Kefalonia Island. The 2014 Kefalonia earthquake sequence (*I1* Figure 3.21) started on 19 January with the first main shock occurring on 26 January ( $M = 6.1$ ), and aftershock activity extending over 35 km (Karakostas et al., 2015), part of which hosted the second main shock ( $M = 6.0$ ) that occurred on 3 February and the compound aftershock activity. A sub-cluster is also de-

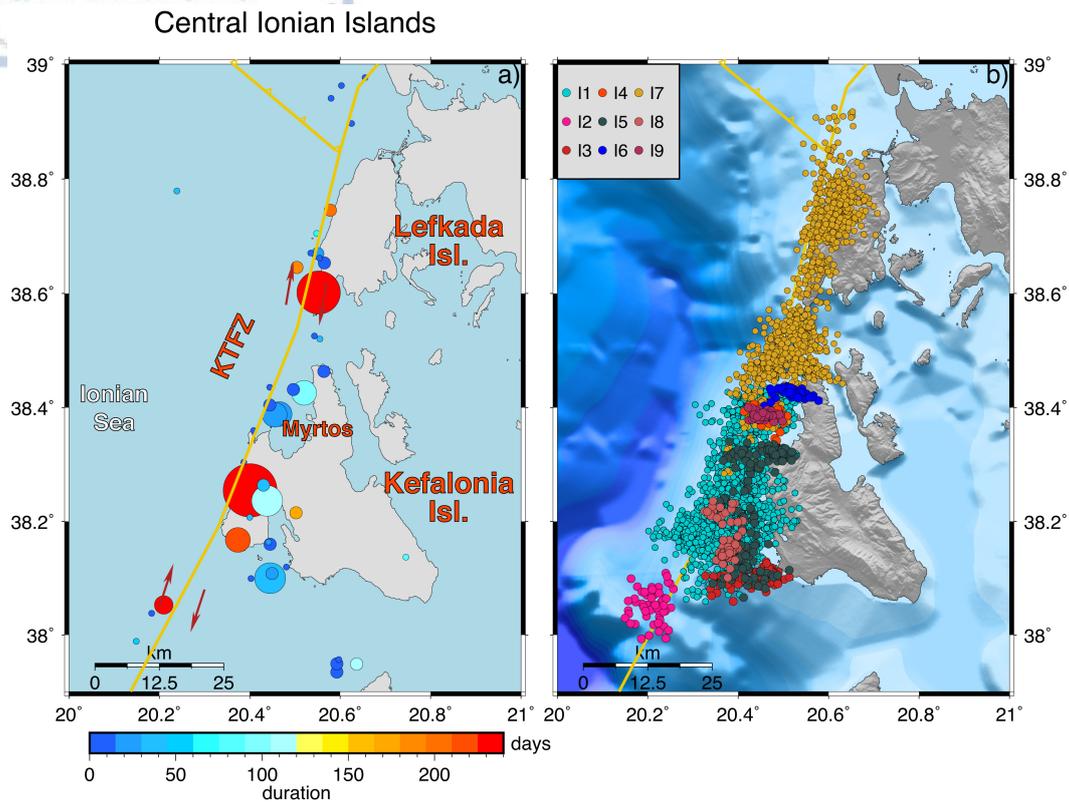


Figure 3.21: (a) Spatial distribution of the centroids of the identified clusters for the area of Central Ionian Islands along with the trace of the Kefalonia Transform Fault Zone (yellow lines). The size of the circles is proportional to the earthquake number in each cluster, whereas the duration is represented by the color scale. (b) Spatial distribution of the clusters with  $N \geq 30$  events. The index of each cluster is given in the inset box.

tected offshore to the southwest of Kefalonia Island ( $I2$  in Figure 3.21b) that is deployed concurrently with the main sequence (Figure C.7). In addition, two distinct clusters,  $I3$  and  $I4$  (Figure 3.21b), are revealed, which occurred between November and December 2014 (Figure C.8), across the edges of the double rupture. They might be triggered by the stress transfer of the main ruptures, indicating activation of adjacent fault segments. The seismic activity of cluster  $I5$  (Figure 3.21b) comprises 164 earthquakes that occurred within 100 days (Figure C.8). It retains the most interest because it is essentially two seismic excitations evolving at the same time. The first initiated in the Myrtos Gulf and the second offshore the south part of Kefalonia Island. The activity

of the *I7* cluster (Figures 3.21b and C.9) spreads along the western coastline of Lefkada and Kefalonia Islands, far beyond both sides of the 2015 Lefkada main rupture. To the south the aftershock activity is sparse, probably due to the large amount of stress released in the main rupture, revealing that the main slip is associated with a fault of about 17 km in length (Papadimitriou et al., 2017). In addition to cluster *I4*, two clusters (*I6* and *I9* in Figure C.8 and Figure C.10, respectively) are detected in the area between Lefkada and Kefalonia, extending to about 15 km, which is considered as a transition zone encompassing step-over structures (Karakostas et al., 2015). All of them relate to the E–W-oriented, parallel step-over faults, similar to the ones detected in the microseismicity cluster analysis between September 2016 and December 2019 in the study area (Bountzlis et al., 2021).

### North Aegean Sea area

The NAS area consists of 187 clusters, including both main shock–aftershock sequences and earthquake swarms (Table 3.14). Figure 3.22a shows that the main clustered activity is concentrated along the NAT and the sub-parallel branches, as well as in the southeastern subarea.

The first seismic excitation with  $N \geq 30$  events (*N1* in Figure 3.22b) is a sequence of interest since two moderate events ( $M = 5.2$  and  $M = 5.3$ ) occurred in 3 weeks, both producing their own aftershocks (Figure C.11). The 2013, January 8  $M = 5.8$  North Aegean earthquake (Karakostas et al., 2014) along with its aftershock activity (cluster *N3* in Figure 3.22b) is also detected. The aftershock activity is temporally divided into two clusters (Figure C.12). The 24 May 2014  $M = 6.9$  Samothraki main shock was followed by aftershock activity confined to three major clusters (*N4*, *N5*, *N6* in Figure 3.22b) and some secondary clusters with  $N \geq 10$  events (Figure C.13), which are in accordance with the ones observed by Saltogianni et al. (2015). The seismic activity that took place near the Aegean coast of NW Turkey during

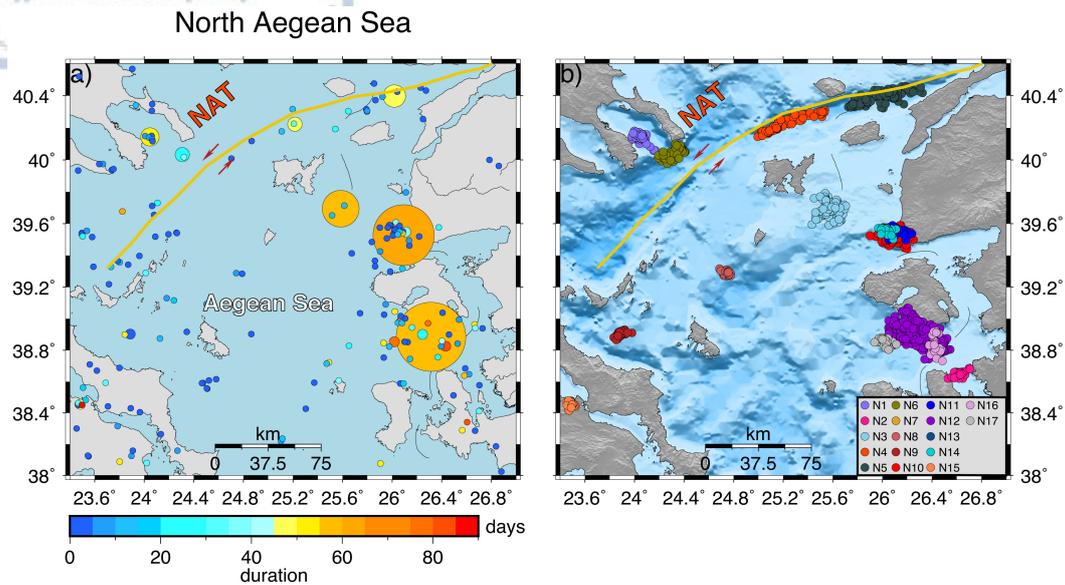


Figure 3.22: (a) Spatial distribution of the centroids of the identified clusters for the area of North Aegean Sea along with the trace of North Aegean Trough (yellow lines). The size of the circles is proportional to the earthquake number in each cluster, whereas the duration is represented by the color scale. (b) Spatial distribution of the clusters with  $N \geq 30$  events. The index of each cluster is given in the inset box.

January–October 2017 (Mesimeri et al., 2018b) is divided into three clusters with  $N \geq 30$  ( $N10$ ,  $N11$  and  $N14$  in Figure 3.22b) and two minor clusters with 22 and 23 events, respectively (Figure C.14). The strong main shock ( $M = 6.4$ ) that occurred on the 12th of June 2017 offshore, south of the SE coast of Lesvos Island, along with its intense aftershock activity, is identified and illustrated in Figure 3.22b ( $N12$ ). Two major ( $N \geq 30$ ) secondary outbursts of clustered activity occurred concurrently on the west ( $N17$ ) and east ( $N16$ ) side of the sequence (Figure C.15). A thorough analysis revealing multiple spatial clusters of the sequences is conducted by Papadimitriou et al. (2018).

### 3.3.2.6 Regional variability of clustering properties

In this section, we investigate regional variations in the clustering behavior of the detected seismic sequences, in particular, on their productivity rates

and on their temporal evolution that can differ among areas with distinct seismotectonic characteristics.

The inverted generic ETAS parameters for the three areas are given in Table 3.15. There are 27 sequences in CG (Figures 3.19 and 3.20), 9 in CII (Figure 3.21) and 17 in NAS (Figure 3.22) from 2012 until 2019 with  $N \geq 30$  events, however, we removed cluster C26 from the computations, since it is located at the boundaries of the study area (Figure 3.20) with part of the aftershock data being omitted.

Table 3.15: Generic ETAS parameter values for the three study areas and the  $\beta$  value of the GR law.  $N^*$  denotes the number of sequences with  $N \geq 30$ .

Area	$p$	$c$	$a$	$K$	$\mu$	$\beta$	$N^*$
CG	1.23	0.0171	0.82	0.74	0.43	2.13	26
CII	1.31	0.11	1.29	0.44	0.15	2.21	9
NAS	1.26	0.0324	1.04	0.51	0.28	2.03	17

The parameter  $a$  for CG ( $a = 0.82$ ) is the lowest among the three areas, indicating the dominance of swarm activity presumably due to fluid flow in accordance with many relevant studies (Mesimeri et al., 2019; Michas et al., 2021). Low  $a$  values characterize areas with high fluid flow (Hainzl and Ogata, 2005), even though the estimated value can be underestimated due to magnitude incompleteness after the occurrence of the main shock or due to the existence of time-dependent background seismicity (Hainzl et al., 2013). Conversely, in CII, the estimated value ( $a = 1.29$ ) is relatively larger compared to the former region ( $a = 0.82$ ), indicating the dominance of typical main shock–aftershock sequences. In the NAS area, a moderate value is acquired ( $a = 1.04$ ), probably due to the co-existence of swarm activity and aftershock sequences. Another indicator for the existence of swarm activity in CG is the large value of the background seismicity ( $\mu = 0.43$ ) compared to NAS and CII. High values of the background rate can indicate the existence of aseismic loading transients (Marsan et al., 2013). Llenos et al. (2009) observed increased values of the background component of the fitted ETAS model when it was

applied to pre-swarm and swarm activity, respectively.

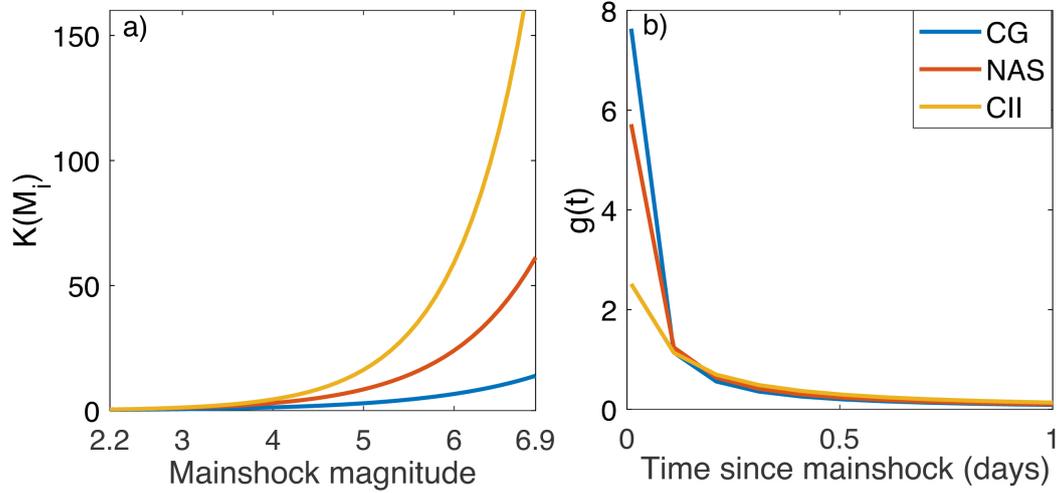


Figure 3.23: (a) The number of events triggered by an earthquake of magnitude  $M_i$  at CG (blue), NAS (orange) and CII (yellow), respectively. (b) The temporal distribution of triggered aftershocks.

Figure 3.23a shows the number of direct triggered events,  $K(M_i)$ , from an earthquake of magnitude  $M_i$  for the three areas. We used Equation 3.7 for the computations and considered  $m_c^* = 2.2$ , which is the maximum completeness magnitude among the three data sets. The exponent  $\beta$  is the mean value over all detected clusters of each area and its value is given in Table 3.15. In CII, the seismic sequences seem to be more productive, as shown in Figure 3.23a, with NAS and CG to exhibit smaller values. Combined with the higher background rate for the area of CG ( $\mu = 0.43$ ), we could say that a significant part of Corinth Gulf's sequences cannot be attributed to the triggering effect of main shocks but different underlying mechanisms seem to play an important role. Conversely, in CII area, main shock–aftershock sequences seem to dominate, generating a rich number of aftershocks (very low background rate,  $\mu = 0.15$ , and high productivity of mother events). Concerning the distribution of aftershocks in time, the normalized Omori law distribution is used, given by Equation A.3 (Figure 3.23b).

### 3.3.2.7. Sequence-specific clustering properties

Next, we estimate the  $a$ -values for the individual sequences of each area by maximizing  $LL$  as a function of  $a$ ,  $K$  and the background rate  $\mu$ , while keeping the rest of the parameters fixed for clusters with  $N < 80$ . In this way, we increase the robustness of the inversion procedure since there are sequences with few events. A similar procedure was followed by [Page et al. \(2016\)](#) and [Llenos and Michael \(2017\)](#) who demonstrated that fitting multiple parameters for a single sequence can be unstable and [Hardebeck et al. \(2019\)](#) who implemented this method for the estimation of California aftershock parameters. We intend to investigate potential differences in the productivity ( $a$ ,  $K$ ) and the background rate,  $\mu$ , among sequences of each area and their relation to different underlying triggering mechanisms. Productivity parameters  $a$  and  $K$  are correlated, so we enabled both to run during the iterative procedure. We also examine the value of the background rate among sequences since it can be also an indicator of aseismic transients in a region. Both parameters,  $a$  and  $K$ , are not influenced by  $\mu$ , as we verified it by implementing the inversion procedure, also keeping parameters  $a$  and  $K$  fixed.

#### Application to Corinth Gulf

In Table 3.16, the inverted parameters for the 26 clusters of dataset D1 with  $N \geq 30$  are given. We adopt the generic values of Omori law ( $p$  and  $c$  in Table 3.15) for clusters with  $N < 80$  to increase the stability of the inverted parameters. We observe relatively high background rates for most of the sequences and low  $a$  values, in particular,  $a < 1$  for 10 out of the 26 clusters.

Concerning the 2013 Aigion earthquake swarm and its subsequent bursts of activity (clusters  $C6$ ,  $C8$  and  $C10$ ), we observe relatively low productivity values of the ETAS model ( $a = 0.20, 0.34, 0.10$ , Table 3.16) in accordance with studies suggesting pore-fluid pressure as the main triggering mechanism

### CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Table 3.16: Details on the 26 clusters with  $N \geq 30$  events in CG area and the inverted ETAS parameters. The generic values of the Omori law,  $p$  and  $c$ , are adopted for clusters with  $N < 80$ .

ID	$T_{in}$	$T_{end}$	$N$	$p$	$c$	$b$	$a$	$K$	$\mu$	$M_{max}$
C1	12/1/12	23/1/12	33	1.23	0.017	1.20	0.49	0.79	1.03	3.1
C2	13/1/12	27/1/12	33	1.23	0.017	0.83	1.69	0.23	1.25	3.1
C3	4/3/12	6/4/12	65	1.23	0.017	1.03	1.53	0.26	1.05	3.0
C4	22/9/12	3/10/12	69	1.23	0.017	0.99	0.36	0.94	1.44	5.0
C5	27/12/12	1/1/13	34	1.23	0.017	0.82	1.84	0.21	1.32	3.8
C6	22/5/13	28/6/13	310	1.45	0.012	0.96	0.20	0.90	0.47	3.7
C7	8/6/13	28/6/13	144	1.11	0.007	1.22	0.60	1.30	1.00	3.0
C8	7/7/13	27/7/13	128	1.04	0.001	0.77	0.34	2.48	0.59	3.7
C9	8/9/13	13/9/13	65	1.23	0.017	1.19	1.28	0.79	2.74	2.8
C10	29/10/13	6/11/13	68	1.23	0.017	1.27	0.10	0.91	2.87	3.1
C11	19/1/14	16/1/14	33	1.23	0.017	0.92	1.26	0.50	1.37	3.8
C12	29/1/14	10/2/14	70	1.23	0.017	0.81	1.39	0.29	1.92	3.9
C13	21/3/14	1/4/14	52	1.23	0.017	0.83	2.97	0.009	3.41	4.0
C14	8/6/14	11/6/14	74	1.23	0.017	0.81	0.92	0.64	4.86	4.3
C15	21/7/14	31/10/14	506	1.37	0.051	1.04	1.38	0.34	1.32	4.6
C16	22/7/14	1/11/14	95	1.26	0.014	1.15	0.72	0.45	0.44	2.8
C17	24/7/14	26/10/14	61	1.23	0.017	0.94	1.72	0.16	0.35	3.4
C18	23/7/14	31/10/14	121	1.25	0.131	0.95	1.77	0.24	0.05	4.7
C19	7/11/14	18/12/14	228	1.07	0.071	0.92	1.80	0.55	0.76	4.8
C20	7/11/14	14/12/14	36	1.23	0.017	1.05	1.27	0.41	0.42	3.1
C21	1/10/15	6/10/15	44	1.23	0.017	1.16	1.97	0.49	1.61	2.8
C22	27/7/16	5/8/16	32	1.23	0.017	0.75	3.50	0.09	0.45	2.7
C23	1/8/16	8/8/16	147	2.79	0.160	0.98	0.10	0.85	2.98	3.4
C24	9/1/17	23/1/17	104	2.79	0.702	0.82	1.70	0.15	1.05	4.5
C25	14/7/17	17/7/17	39	1.23	0.017	0.43	0.73	0.40	5.95	4.2
C27	30/10/17	2/11/17	31	1.23	0.017	0.50	1.68	0.10	6.19	3.5

during the excitation (Kapetanidis et al., 2015). Clusters  $C11$  and  $C12$  are part of the same swarm (Figure C.3) that occurred offshore Psathopyrgos fault. Their relatively high background rates ( $\mu = 1.34, 1.92$ ) show that a significant part of the clustered seismicity cannot be explained by the empirical laws of the triggering part of the ETAS model. Cluster 14 is part of a major swarm that began on 8 June 2014 (Figure C.3). Michas et al. (2021) did not find high diffusion rates that are related to fluid pore pressure. However, the large background rate found in our study ( $\mu = 4.86$ ) and the low  $a$  value ( $a = 0.92$ ) suggest the existence of a non-typical main shock–aftershock sequence, with more complex triggering mechanisms being responsible, such as aseismic creep.

## CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Similarly, the largest cluster in the dataset, the  $C15$ , located offshore Nafpaktos, is characterized by relatively high background rate ( $\mu = 1.32$ ) and low productivity ( $a = 1.38$ ), more typical values for swarm activity. In contrast, clusters  $C18$  and  $C19$  that are more typical main shock–aftershock sequences with a distinct in magnitude event in the initiation of the sequence (Figure C.4), have low background rates ( $\mu = 0.05, 0.76$ ) and relatively high productivity rates ( $a = 1.77, 1.80$ ). The two clusters near Itea Gulf show contradictory results and in particular, the first one,  $C13$ , is characterized by a high background rate ( $\mu = 3.41$ ), whereas the second,  $C17$ , which occurred four months later, exhibits a much smaller background value ( $\mu = 0.35$ ) more typical for main shock–aftershock sequences. However, biases can exist in the inversion of the parameters for clusters with a small number of events, so we should be cautious with the characterization and interpretation.

### Application to Central Ionian Islands

In Table 3.17, the inverted parameters for the nine clusters identified in the area of CII with  $N \geq 30$  events are given, where we kept fixed the Omori law parameters  $p$  and  $c$  (generic values in Table 3.15) for clusters with  $N < 80$  during the estimation procedure.

Table 3.17: Details on the 9 clusters with  $N \geq 30$  events in CII area and the inverted ETAS parameters. The generic values of the Omori law,  $p$  and  $c$ , are adopted for clusters with  $N < 80$ .

ID	$T_{in}$	$T_{end}$	$N$	$p$	$c$	$b$	$a$	$K$	$\mu$	$M_{max}$
I1	19/1/14	16/9/14	2829	1.42	0.24	0.79	1.31	0.40	0.17	6.1
I2	23/1/14	14/9/14	55	1.31	0.11	1.23	1.38	0.30	0.12	3.7
I3	5/11/14	11/12/14	134	1.36	0.06	0.99	1.44	0.29	0.99	5.1
I4	13/11/14	12/12/14	66	1.31	0.11	0.93	1.43	0.38	0.37	4.9
I5	5/1/15	27/4/15	164	1.05	0.01	0.93	2.82	0.10	0.76	4.4
I6	18/1/15	24/4/15	71	1.31	0.11	1.08	1.91	0.36	0.15	3.8
I7	13/11/15	26/6/16	1396	1.45	0.30	0.86	1.51	0.29	0.45	6.5
I8	20/11/15	25/6/16	65	1.31	0.11	0.84	0.94	0.53	0.07	4.3
I9	4/4/17	4/5/17	67	1.31	0.11	0.95	2.26	0.18	0.70	3.9

The estimated ETAS parameters of the sequence  $I1$  are in accordance with the existence of a main shock–aftershock sequence described in Section 3.3.2.5. In particular, the background rate is relatively low ( $\mu = 0.17$ ), indicating that the seismicity is adequately described by the triggering part of the ETAS intensity function. The seismic activity of clusters  $I3$  and  $I5$  (shown by green and blue color in Figure C.8, respectively) are characterized by relatively high background rates ( $\mu = 0.99, 0.76$ , Table 3.17). The space-time evolution of the former indicates a rapid migration in the beginning of the sequence (Figure C.8), whereas, for the latter, it is characterized by the smallest  $K$  value ( $K = 0.10$ ) in the area although the  $a$  value is rather large. Taking into account the lack of distinct main shocks at the initiations of the sequences, they can be characterized as earthquake swarms, one of the few observed in an area which comprises mostly main shock–aftershock sequences. Concerning cluster  $I9$ , located in the transition zone between Lefkada and Kefalonia Islands, there is evidence for swarm activity due to the relatively high background seismicity rate ( $\mu = 0.70$ ). Ultimately, the major main shock–aftershock sequences in the area,  $I1, I7$ , get the highest  $p$  values ( $p = 1.42, 1.45$ ), meaning that they are characterized by rapid aftershock decay in time.

### Application to North Aegean Sea

In NAS area cluster  $N1$ , which comprises two moderate events ( $M = 5.2$  and  $M = 5.3$ ) within a period of 3 weeks, exhibits the lowest  $a$  value ( $a = 1.10$ ) among the main detected clusters, which could be an indicator of fluid diffusion in the area (Table 3.18). Another case worth mentioning is the 24 May 2014,  $M = 6.9$ , Samothraki seismic sequence which is divided into three major clusters ( $N4, N5, N6$ , in Figure 3.22). The estimated background rates of the three major clusters are relatively small ( $\mu = 0.16, 0.60, 0.29$ ), whereas the opposite holds for the scaling parameter,  $a$ , for the first two clusters ( $a = 1.82, 1.76$ ). Concerning the seismic excitation that consists of clus-

### CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

Table 3.18: Details on the 17 clusters with  $N \geq 30$  events in NAS area and the inverted ETAS parameters. The generic values of the Omori law,  $p$  and  $c$ , are adopted for clusters with  $N < 80$ .

ID	$T_{in}$	$T_{end}$	$N$	$p$	$c$	$b$	$a$	$K$	$\mu$	$M_{max}$
N1	14/2/12	4/4/12	136	1.41	0.03	0.96	1.10	0.40	0.54	5.3
N2	27/4/12	3/5/12	30	1.26	0.03	0.53	1.64	0.16	1.07	4.8
N3	8/1/13	6/3/13	285	1.07	0.06	0.88	2.39	0.06	0.55	5.8
N4	24/5/14	9/7/14	94	1.41	0.78	0.74	1.82	0.01	0.16	6.9
N5	24/5/14	11/7/14	153	1.60	0.16	0.69	1.76	0.16	0.60	4.5
N6	24/5/14	22/6/14	83	1.49	0.04	0.64	1.25	0.30	0.29	4.4
N7	6/12/14	29/12/14	41	1.26	0.03	0.67	1.60	0.15	0.31	4.9
N8	26/3/15	2/4/15	30	1.26	0.03	0.97	1.45	0.36	1.76	4.1
N9	29/10/16	31/10/16	49	1.26	0.03	0.89	2.44	0.28	2.88	3.4
N10	26/1/17	28/3/17	568	1.29	0.04	0.73	1.31	0.36	1.00	5.1
N11	7/4/17	12/5/17	38	1.26	0.03	1.05	1.29	0.11	0.91	3.4
N12	12/6/17	8/8/17	614	1.48	0.12	0.79	1.46	0.25	0.86	6.4
N13	13/6/17	29/7/17	48	1.26	0.03	1.03	2.42	0.17	0.35	3.7
N14	15/8/17	23/10/17	38	1.26	0.03	1.06	1.13	0.39	0.26	3.5
N15	16/8/17	11/11/17	34	1.26	0.03	1.08	1.46	0.36	0.15	3.5
N16	17/8/17	8/11/17	39	1.26	0.03	1.24	2.39	0.14	0.31	3.2
N17	24/8/17	11/11/17	35	1.26	0.03	1.01	2.23	0.13	0.27	3.6

ters  $N10$ ,  $N11$  and  $N14$ , the relatively low productivity rates of the ETAS model ( $a = 1.31, 1.29, 1.13$ ) and, conversely, the relatively high background rates for the first two,  $N10$  and  $N11$ , clusters ( $\mu = 1.00, 0.91$ ) may indicate fluid intrusion. This observation is in accordance with the study of [Mesimeri et al. \(2018b\)](#) who derived high background rates after the estimation of the ETAS model to the empirically divided 5 sub-clusters of the primary seismic activity (January–March 2017). A fast-diminishing aftershock activity is observed for the main shock ( $M = 6.4$ ) that is located SE of Lesvos Island ( $N12$ ), which is translated into a high Omori exponent,  $p = 1.48$ . Additionally, low background rates characterize the three main clusters,  $N12$ ,  $N16$  and  $N17$ , indicating that they are probably related to tectonic and coseismic stress transfer from previous seismicity ([Papadimitriou et al., 2018](#)). Worth mentioning are the remarkable high background rates for clusters  $N8$  ( $\mu = 1.76$ ) and  $N9$  ( $\mu = 2.78$ ), which could be an indicator for seismic activity driven by transient forces, however, the number of events is rather small and could have

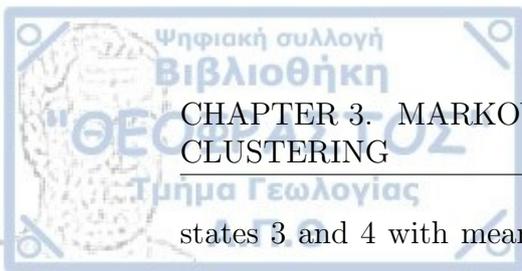
led to significant biases in the inversion of the parameters.

### 3.4 Summary and discussion

In the first part, we introduce the MAP model and its use as a change point tool for seismicity rates. We evaluated its ability to detect changes in the seismicity rate related either to main shock-aftershocks or earthquake swarms to the 1964-2017 Corinth Gulf earthquake catalog with  $M \geq 4.5$ . This is a data set that includes well studied seismic sequences that facilitate the evaluation of the model.

Concerning the estimation of the parameters we used the grid-based method for the choice of the initial values of  $EM$  algorithm and the parallel framework (described in Section 2.4) reducing the highly demanding computation cost. AIC and BIC demonstrate that the temporal evolution of the 1964-2017 earthquakes with  $M \geq 4.5$  in the two subareas of Corinth Gulf can be described through the adoption of four occurrence rates, i.e., low, moderate, high and very high and the duration of the associated periods can be estimated by the model. The residual analysis shows that the MAP model captures the main temporal features of the observed seismicity. In addition, we compared the MAP model with the MMPP which belongs to the same class of stochastic models and it has been used for the modeling of earthquake's temporal distribution (Benali et al., 2020; Lu and Vere-Jones, 2011). Their difference lies on the structure of their rate matrices. In MAP, changes in the seismicity rate can be triggered by a previous event, while under the MMPP formulation seismicity rate can change multiple times before the earthquake occurrence. We believe that these conditions have an impact on the adaptability of the model to the data, as the MAP model shows better results compared to MMPP in terms of AIC and BIC differences.

The evaluation of the transitions among the hidden states of the model showed that main shocks with  $M \geq 5.6$  in the western subarea occurred in



### CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

---

states 3 and 4 with mean sojourn time until the next earthquake occurrence 9 hours and 11 minutes, respectively. This means that an aftershock with  $M \geq 4.5$  is expected in the next hours after the occurrence of an earthquake with  $M \geq 5.6$ . States 3 and 4 correspond mainly to main shocks and the immediate aftershocks in both subareas, whereas mostly in the western subarea state 2 in some cases corresponds to secondary aftershocks and foreshocks. In the eastern subarea a more episodic behavior dominates, since intense seismic periods of short duration alternate with long relatively quiescent ones.

Since our model is purely temporal, it would be challenging to apply the MAP on induced seismicity catalogs, where the spatial area is restricted, in order to detect changes in the seismicity rates. Assigning a spatial component could increase the association between seismicity rate changes and seismic excitation, as there are cases where events occur very close in time but not in space.

The two-step clustering procedure, MAP-DBSCAN, for the detection of spatiotemporal seismic clusters combines the Markovian Arrival Process for an initial separation of the background seismicity from potential seismic excitations detecting changes in the seismicity rate and a density-based clustering algorithm, DBSCAN, for the detection of areas with high spatial density. The consistency and efficiency of the MAP-DBSCAN method is examined on a simulated ETAS catalog that produces the main features of seismicity in the region of Greece. In particular, we showed that our method is able to identify the connections among the events generated by a spatiotemporal ETAS model, as well as the mother events that initiated each cluster. The knowledge of the links among the events enabled the comparison of the method with some well known clustering algorithms, like the Gardner and Knopoff, the Reasenberg and the Nearest-Neighbor, by the use of the Jaccard index. This is a tool for measuring the overlap between the original partition of events into clusters and background seismicity, and the estimated one after the implementation of each clustering method. The results show that MAP-DBSCAN method is very com-

petitive and in most cases outperforms the tested algorithms. The NN achieves the best reconstruction of the clusters (Table 3.12), which is probably related to the similarity of its metric with the ETAS one that is used for the generation of the seismicity. The window-based method overestimates the clustered seismicity in accordance with the observations made by [Peresan and Gentili \(2020\)](#), whereas the Reasenber link-based method seems to overestimate the background events (Figure 3.14).

We applied the method to three major seismic zones in Greece during 2012–2019, identifying the major seismic sequences and a plethora of smaller ones which are in accordance with the main seismotectonic properties of the study areas. Depending on the quality of the data set the method can be also used for the detection of secondary faults ([Bountzis et al., 2021](#)). The rich seismic activity during 2013-2014 in the western subarea of the Corinth Gulf is detected in detail, a nontrivial issue, especially for the area between Nafpaktos-Psathopyrgos and offshore Aigion, where multiple excitations occurred in close proximity and within short periods (Figures 3.19, C.4 and C.5). Seismicity in the eastern subarea of the Corinth Gulf is found to be more sparse with few major clusters located near Itea Gulf (Figures 3.20 and C.7) and offshore Perachora and Xylokastro (Figure 3.19 and C.6). On the contrary, seismicity in the Central Ionian Islands is dominated by the 2014 Kefalonia and the 2015 Lefkada seismic sequences (Figure 3.21). Together they comprise the 81% of the clustered seismicity in this area. Many large clusters are identified in the North Aegean Sea area that includes both main shock–aftershock sequences and earthquake swarms.

We investigated the properties of clustering seismicity among the three study areas with the use of the ETAS model. The results indicate that there are differences in aftershock productivity rates between Corinth Gulf, Central Ionian Islands and North Aegean Sea, showing that productivity can vary regionally. As showed by [Page et al. \(2016\)](#) and [Llenos and Michael \(2017\)](#) adopting the regional variations of productivity can produce a significant gain

on aftershock forecasts. In the Central Ionian Islands, main shock–aftershock sequences seem to be more productive with the North Aegean Sea and the Corinth Gulf to follow (Figure 3.23). The sequences in the Corinth Gulf in particular are characterized by the highest background rate among the three areas (Table 3.15), meaning that a significant portion of clustered seismicity is not caused by the triggering of a main shock coseismic slip, but by the contribution of different triggering mechanisms. Many studies have focused on this area, suggesting pore-pressure changes due to fluid migration and aseismic creep as possible triggering mechanisms for the clustered seismicity (Kapetanidis et al., 2021; Mesimeri and Karakostas, 2018). In the North Aegean Sea, the swarm activity coexists with aftershock sequences, implying that for forecasting purposes, a finer regionalization might be more appropriate.

We also investigated potential differences in the productivity and the background rates among sequences of each region and their relation to different underlying triggering mechanisms. Results show that the high background seismicity ( $\mu$ ) and low productivity ( $a$ ) values of the ETAS model are related to earthquake swarm activity triggered by fluid pore-pressure changes, such as the 2013 Aigion swarm (clusters  $C6$ ,  $C9$  and  $C10$ , Table 3.16, Figures 3.19 and C.3) in Corinth Gulf (Kapetanidis et al., 2015) and the 2017 Tuzla earthquake swarm (clusters  $N10$ ,  $N11$  and  $N14$ , Table 3.18, Figures 3.22 and C.15) in North Aegean Sea (Mesimeri et al., 2018b). This is in accordance with studies suggesting the dependence of low productivity values to the existence of fluids (Hainzl and Ogata, 2005; Hainzl et al., 2013). In general, 18 out of 26 clusters in Corinth Gulf have background rates  $\mu > 1$  and low productivity values (11 out of 26 with  $a < 1$ ), whereas in the Central Ionian Islands, where main shock–aftershock sequences dominate, we observe very low background rates of the ETAS model (all with  $\mu < 1$ ) and relatively high productivity values. In the North Aegean Sea area, we cannot observe a clear pattern, however, the majority of the detected clusters are characterized by low background rates and relatively high productivity, suggesting the dominance of

typical main shock–aftershock sequences.

The advantage of using the MAP model lies in its generality, as any stationary point process can be approximated by a sequence of MAPs under the framework of multiple embedded seismicity rates ([Asmussen and Koole, 1993](#)). It can be extremely beneficial in detecting different types of clusters (dynamic modeling of seismicity rate) by capturing the changes in seismicity rate, without considering any physical assumption, in the sense that each hidden state or group of states might be associated to a different underlying triggering mechanism. Furthermore, in case of non-stationary background seismicity, the MAP model can approximate the different phases by embedding multiple states into the Markov process  $J_t$ , i.e., distinct occurrence rates, and adopting a multiple rate threshold alternating according to the phase of the process at each time. In this way, although it is more complicated, we can model both the non-stationary background seismicity and the triggered events without declustering the earthquake catalog ([Bountzlis et al., 2021](#)). For instance, [Benali et al. \(2020\)](#) applied a Markov Modulated Poisson Process for the description of the background seismicity in northeastern Italy after removing all the triggered events, suggesting the existence of three distinct average trends. The ETAS model assumes a stationary background seismicity rate that is sensitive to transient aseismic forces such as fluid intrusion, and leads to poor fitting results on data sets that include seismic sequences of swarm type according to [Kumazawa and Ogata \(2014\)](#). [Lombardi and Marzocchi \(2010\)](#) showed the inadequacy of a stationary ETAS model with constant background rate to reproduce the temporal patterns of observed seismicity in the Umbria-Marche region due to the increase of the background seismicity rate after the repeated Colfiorito main shocks, as a consequence of the perturbation to the coseismic stress field. They approximated the non-stationary background rate by fitting stationary ETAS models to data in moving windows. Moreover, the modified Omori law which is used to model the aftershock activity, is difficult to predict how long an earthquake swarm may last. The DBSCAN algorithm does not



assume any specific spatial distribution of earthquakes and settles them into groups based solely on their spatial density.

### 3.5 Conclusions

We proposed a novel stochastic modeling for earthquake's temporal distribution through Markovian Arrival Process and the use of a local-decoding algorithm for the detection of changes in seismicity rate. The application of the model to the Corinth Gulf earthquake catalog during 1964-2017 for events with  $M \geq 4.5$  showed that the model efficiently captures the evolution of seismicity in time and shows better results than the Markov Modulated Poisson Process, a model that belongs to the same class of Markov processes. We verified on well studied seismic sequences that the model is capable to detect changes in the seismicity rate related either to main shock-aftershocks or earthquake swarms, so it can be used as a change point tool for seismicity rates.

We established a new clustering procedure, MAP-DBSCAN, that is a combination of the MAP model for an initial separation of the background seismicity from potential seismic excitations, using the detected changes in the seismicity rate, and the density-based clustering algorithm, DBSCAN, for the detection of elevated spatial density areas. We demonstrate the efficiency of the clustering procedure on a simulated ETAS earthquake catalog where the structure of the clusters is known a priori. We proposed the Jaccard index as a validation metric, which we believe is an appropriate tool that incorporates not only the correct links but also the false and missed ones in its form and we encourage its use in performance studies for earthquake clustering. Additionally, we introduced a similar metric for the validation of the model to identify the mother events of each cluster. Moreover, we showed the competitiveness of the MAP-DBSCAN procedure against well-known clustering algorithms, as in most cases, exhibits better results. The method is applied on earthquake catalogs of three major seismic zones in Greece and their clustering properties



### CHAPTER 3. MARKOVIAN ARRIVAL PROCESS FOR EARTHQUAKE CLUSTERING

---

are investigated with the use of the temporal ETAS model. The main seismic clusters in the Corinth Gulf, Central Ionian Islands and North Aegean Sea during 2012–2019 for events with  $M \geq 1.5$ ,  $M \geq 2.2$  and  $M \geq 2.1$ , respectively, are detected by our method, revealing the main seismotectonic structures of the areas.

A stacking procedure is implemented for inverting the generic ETAS parameters of each area. In particular, a common log-likelihood is defined that is the product of the log-likelihood of each individual identified cluster with  $N \geq 30$ . The inverted parameters correspond to the maximization of the product. Regional variability in aftershock productivity and background rates among the areas is observed. The Corinth Gulf is characterized by low productivity values and high background rates related to the dominance of earthquake swarms, whereas seismicity in the Central Ionian Islands consists of main shock–aftershock sequences with high productivity.

We also inferred sequence-specific parameters of the temporal ETAS model implementing a parallel procedure for reducing the computational cost. We evidence the dependence between low productivity values and high background rates with pore-pressure due to fluids migration. We believe that future studies on Operational Earthquake Forecasting should incorporate localized parameters into the models to improve the forecasting accuracy.





## Chapter 4

# Markovian Arrival Process for forecasting large earthquakes number

### 4.1 Introduction

The apparent increase in the occurrence frequency of the great ( $M \geq 8.0$ ) global earthquakes since 2004 (Beroza, 2012), led many authors to investigate the existence of non-stationarity in their temporal distribution. Lay (2015) observed an increase in the frequency of great subduction earthquakes ( $M \geq 8.0$ ) in circum-Pacific belt after 2004. Bufe and Perkins (2005) suggested the existence of mega-quakes ( $M \geq 8.6$ ) clustering based on the apparent concentration of events during the time period 1950-1965 which was followed by 36 years of seismic quiescence. They generated 100,000 simulations of Poisson simulated catalogs to test the times the apparent clustering during 1950-1965 is observed. They found that only in 2% of the catalogs this behavior can be repeated by chance. More recently, Luginbuhl et al. (2018) investigated whether the temporal distribution of large global earthquakes ( $M \geq 7.0$ ) during 2004-2016 departs from Poisson process and exhibits clustering. They

found statistical significant deviations from Exponential distribution for the inter-event times based on a large number of synthetic catalogs. There are also regional studies that examine the existence of time-dependency and non-stationarity for large earthquakes. [Faenza et al. \(2003\)](#) evidenced long-term changes in the seismicity rate of earthquakes with  $M \geq 5.0$  since 1600 in Italy, and [Iliopoulos et al. \(2020\)](#) found long-range memory for earthquakes with  $M \geq 6.5$  in Greece during 1845-2017.

The brevity of the seismological record, along with the limited number of global earthquakes influence the robustness of the statistical tests to identify non randomness in their occurrence ([Daub et al., 2015](#); [Dimer de Oliveira, 2012](#)) leading to a large debate in the scientific community with contradictory results ([Kerr, 2011](#)). However, we believe that even if the hypothesis of the stationary Poisson process for the temporal distribution of global earthquakes cannot be rejected with high statistical power, we should not grant as negligible the possibility for a better approximation of the physical process with non-stationary stochastic models, that could combine the long-term properties of seismicity with the short-term clustering of events. Towards this direction we establish a two-step modeling procedure of the Markovian Arrival Process to model the temporal distribution of long quiescence periods and short-term seismicity simultaneously that exist in large earthquake catalogs and we implement catalog-based pseudo-prospective forecasting experiments for the full distribution of the occurrence frequency to evaluate its performance.

An important assumption concerning the features of the large earthquakes temporal patterns is the existence of long-term changes in the occurrence rate, between extended periods of seismic quiescence with long inter-event times that characterize the tail of their distribution and periods of moderate seismic activity. We show that their temporal behavior cannot be captured well by the MAP model due to the presence of short inter-event times. The strong spatiotemporal concentration of seismicity often obscures long-term features that may characterize the earthquakes temporal distribution ([Zaliapin and](#)

## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

---

[Ben-Zion, 2022](#)). In our study, we will model both temporal behaviors of the large earthquakes distribution by considering the long inter-event times as extreme events due to their rarity. Statistical analysis and forecasting in problems that incorporate extreme events is known to be highly complex as the short times do not conform well with the rare large values ([L'vov et al., 2001](#)). Based on that, we proceed with a modification on the formulation and propose a two-step estimation procedure of the model, where the extreme events are estimated separately from the short-time values. This procedure has been applied to assess the volcanic hazard for the Canary Islands by [Sobradelo et al. \(2011\)](#). They considered time-dependence for the series of eruptions and used a non-homogeneous Poisson process to forecast future events. This method can also be seen in environmental ([Eastoe and Tawn, 2009](#); [Towe et al., 2020](#)) and financial ([Schneider et al., 2021](#)) applications among others, where the extremes are estimated separately under the Extreme Value Theory, ignoring the potential effects of the short-time data. An important parameter that needs to be defined is the threshold value above which events are considered extremes ([Hill, 1975](#); [Smith, 1989](#)).

The aim of this approach is to assess if the MAP model with non-stationary characteristics contributes to the forecasting of the large earthquakes number. For that purpose, we present two main approaches. Firstly, we introduce a two-step modeling procedure based on the extreme values of the observations. Then, we implement a pseudo-prospective experiment based on simulations of the earthquake temporal distribution for the comparison against the Poisson, non-Poissonian renewal models and the temporal ETAS model.

## 4.2 π. Data

### 4.2.1 Catalog description

We focus on megathrust environments where the amount of data permits a robust statistical analysis on the temporal features of large earthquake occurrences. We considered an earthquake catalog comprising subduction earthquakes, trenches, outer rise and overriding plate earthquakes (Figure 4.1) in the circum-Pacific belt. The data set is taken from the International Seismological Centre–Global Earthquake Model (ISC–GEM) version 7.0 of the worldwide earthquake catalog (Di Giacomo et al., 2018; Storchak et al., 2013, 2015), supplemented by the U.S. Geological Survey (USGS) catalog (<http://earthquake.usgs.gov/earthquakes>), up to 31 December 2020.

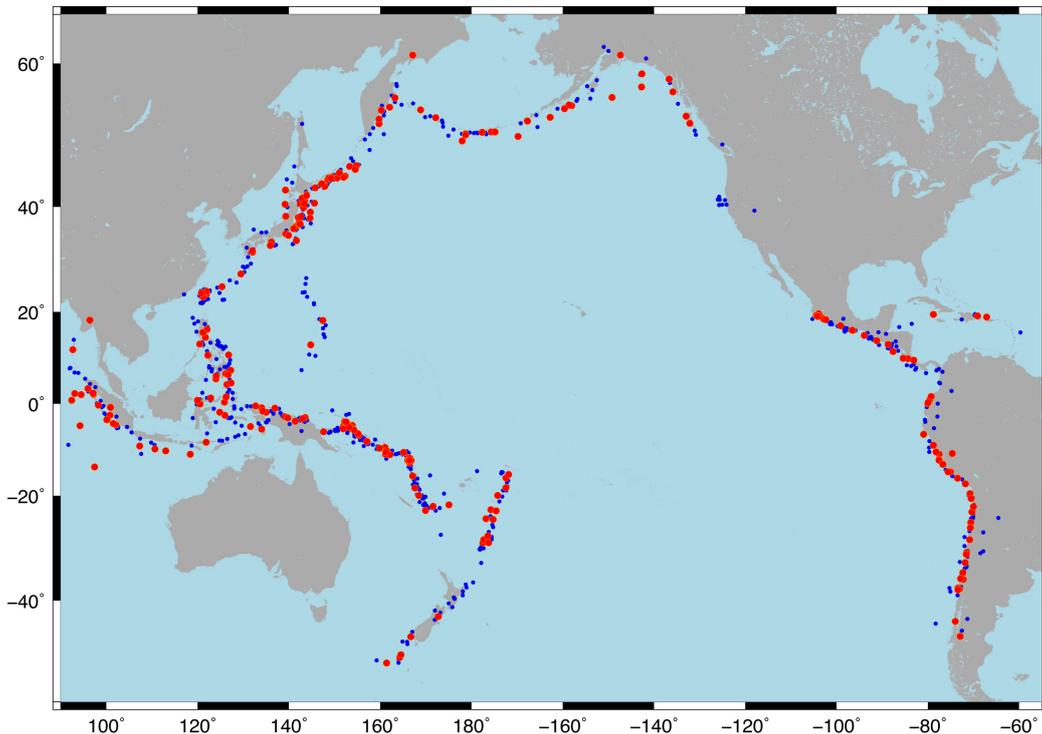


Figure 4.1: Epicentral map of the 909 earthquakes with  $M \geq 7.0$  that occurred from 01 January 1918 up to 31 December 2020 in the circum-Pacific belt. Events with  $7.0 \leq M < 7.6$  and  $M \geq 7.6$  are shown in blue and red circles, respectively.

The magnitude of completeness for shallow earthquakes of the ISC-GEM earthquake catalog is found equal to  $M_c = 7.0$  since 1918 (Michael, 2014). Taking this into account, we applied the MAP model to an earthquake catalog with the magnitude cutoff equal to the completeness threshold, i.e.,  $M_{thr} = 7.0$  including 909 events and three data subsets with corresponding magnitude thresholds  $M_{thr} = 7.6$ ,  $7.7$  and  $M_{thr} = 7.8$  that contain 239, 178, 138 events since 1918, respectively. We consider different magnitude cutoffs to investigate the effect of the minimum magnitude,  $M_{thr}$ , on the forecasting gain of the MAP compared to other statistical models. Essentially, we want to examine whether the non-stationarity that we consider through the MAP modeling for the temporal occurrence of large earthquakes is a universal characteristic or it depends on the adopted magnitude range. The maximum focal depth is set to 60 km, which is a reasonable lower cutoff of interplate subduction zone earthquakes according to Scholz (2019).

An issue that is encountered in studies that investigate the existence of interactions among large earthquakes, for instance on whether the surge of great earthquakes ( $M \geq 8.0$ ) since 2004 is random or not, is the use of a declustering algorithm for the separation of triggered from independent seismicity. Since these algorithms tend to produce declustered catalogs with stable rate (Poissonian) it could have a direct impact on our investigation. Moreover, we believe that each earthquake is considered distinct, with its own tectonic context and failure process, comprising equally important threat in terms of hazard. Finally, removing potential aftershocks and foreshocks would reduce the size of the data sets, decreasing the statistical power of the goodness of fit tests. Hence, in this study we consider the complete earthquake catalogs without removing short-term triggered seismicity.

### 4.2.2. Non-stationary features of the catalog

Figure 4.2 illustrates summary results for the four data sets, where similar patterns are exhibited independently of the magnitude thresholds. The seismic moment (in  $Nm$ ) is calculated from the reported magnitude (all magnitudes in ISC-GEM and USGS catalogs are reported as  $M_w$ ) according to the relation  $M_0 = 10^{1.5M_w+9.05}$  (Hanks and Kanamori, 1979).

From the inter-event times distribution in Figure 4.2a we can visually recognize a distinct peak of small values (black vertical arrow) for all magnitude thresholds that could be attributed to localized aftershocks and foreshocks, as well as two groups of events with different sets of values (horizontal dashed dot bidirectional arrows), although less pronounced for  $M \geq 7.0$ . The first group can be associated to periods of relatively high seismic activity (red color) compared to the second one (blue color) that consists of some quite large values corresponding to long periods of relative seismic quiescence such as in the mid-1920, mid-1950 and early 1980 (Figure 4.2c). This could be an indicator that the inter-event times distribution has heavy-tailed characteristics, i.e., the values in the tail are not bounded by an exponential distribution. Finally, the abrupt jumps of the cumulative seismic moment in 1960 and after 2004 (Figure 4.2b) coincide with periods of increased seismicity. This agrees with Zaliapin and Kreemer (2017) who observed changes in the moment release during these periods.

We argue that the temporal characteristics of the earthquakes show at least after visual inspection (Figure 4.2), that seismicity can be temporally divided into three distinct periods. For instance, there are periods of relative seismic quiescence alternating with periods of relatively high seismicity. Specifically, in 1924-1927 2 events ( $M \geq 7.6$ ) occurred then a 5-year period, 1928-1932, with 14 events ( $M \geq 7.6$ ) follows. In 1973-1978 14 events ( $M \geq 7.6$ ) occurred and then a 6-year period, 1979-1984, follows with 6 events ( $M \geq 7.6$ ). The third period exhibits more intense seismicity related to foreshocks and

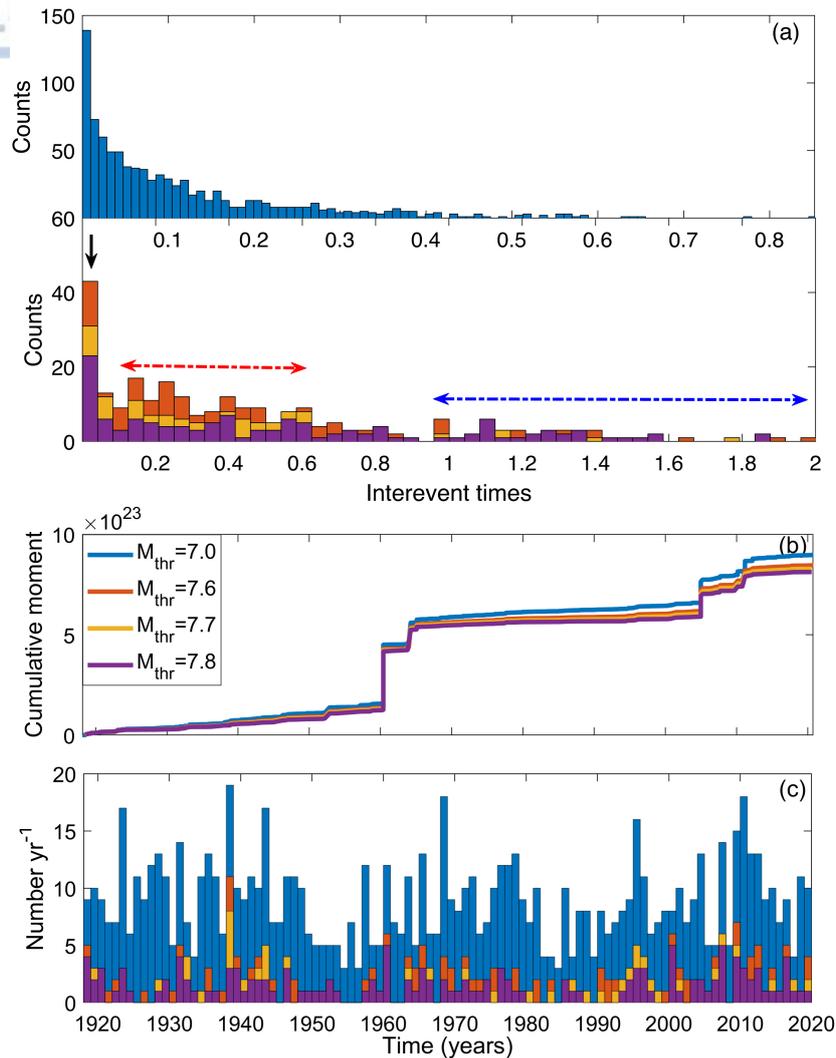


Figure 4.2: Temporal characteristics for  $M \geq 7.0$  (blue),  $M \geq 7.6$  (orange),  $M \geq 7.7$  (yellow) and  $M \geq 7.8$  (purple) earthquakes that occurred from 01 January 1918 to 31 December 2020 in the circum-Pacific belt. (a) Inter-event time distribution. A peak of short inter-event times is evident (black vertical arrow), probably related to localized aftershocks and two groups of inter-event times, indicated by the dashed-dot horizontal bidirectional arrows, that can be interpreted as periods of increased seismicity (red color) and of relative seismic quiescence (blue color), respectively. (b) Cumulative seismic moment (in  $Nm$ ) release as a function of time. (c) Yearly seismicity rate as a function of time.

aftershocks, like in 1960 with the giant ( $M_w 9.6$ ) Chilean earthquake and its two foreshocks ( $M_w = 8.1, 8.6$ ) within 2 days or in 1938 with the four large ( $M_w = 7.8, 7.7, 7.7$  and  $7.6$ ) Honshu earthquakes.

The yearly seismicity rate in Figure 4.2c indicates the existence of idle

periods, namely periods with low seismicity rate and consequently longer inter-event times for the large earthquakes. Additionally, the boxplots of the inter-event times in Figure 4.3 show that events outside the boxes appear in all data sets, indicating variability above the upper quantiles. These long inter-event times can be seen as extreme events rather than outliers, which characterize the tail of the inter-event times distribution.

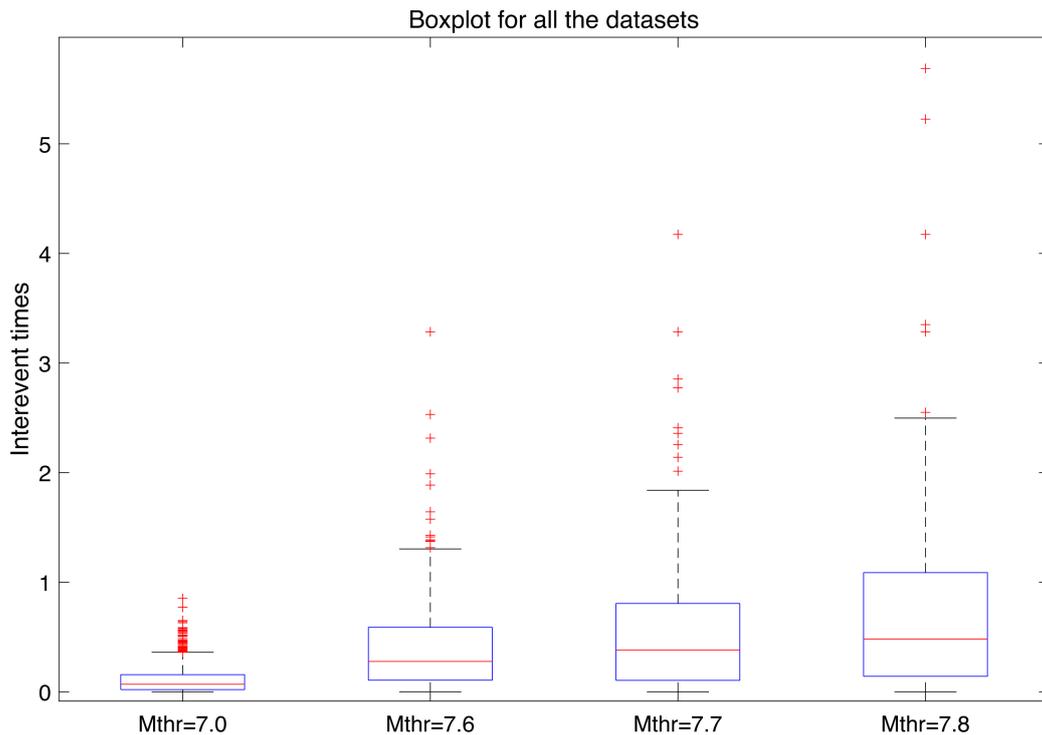


Figure 4.3: Boxplots for the four datasets tested in this study. In all cases, there is a significant subset of events outside the third quartile plus one standard deviation.

### 4.3 Methodology

The framework of a hidden process whose states modulate the seismicity rate,  $\lambda_{J_t}$ , of a counting process,  $N_t$ , as in MAPs, is anticipated to facilitate the modeling of changes in the earthquake occurrence rate when the responsible physical mechanisms are unknown. The occurrence rate of the MAP model

is varying according to the transitions among the states of the latent Markov process,  $J_t$ , making it suitable for the investigation of the assumption of non-stationarity for the seismicity rate of large earthquakes. The forecasting probabilities in the MAP framework depend on the current state of the process  $J_t$ , for instance, when the model is in a state with high occurrence rate (interpreted as a period of increased seismicity) the probability of an earthquake occurrence is different (higher) than the one when the model is in a state with low occurrence rate (period of relative seismic quiescence). Hence, we compare the forecasting performance of the proposed MAP model described in Chapter 2 with other renewal models as well as with the standard ETAS model. Better performance can indicate the significance of embedding non-stationary characteristics to the modeling procedure.

### 4.3.1 Establishment of MAP model with an "idle" state

Our main assumption is that long-term changes in the seismicity rate might exist, where quiescence periods alternate with periods of moderate seismic activity. However, the long-term features of the large earthquakes cannot be captured simultaneously with the short inter-event times. Hence, we proceed with a modification on the formulation of the model and propose a two-step estimation procedure where the long inter-event times that characterize the tail of the temporal distribution can be seen as extreme events due to their rarity and their occurrence rate will be estimated separately from the rest of the data set.

Firstly, we consider the existence of an "idle" state that corresponds to periods of relative seismic quiescence with long inter-event times, where events occur according to a Poisson process with occurrence rate,  $\lambda_{idle}$ . Let  $Tr = \{\tau_1, \dots, \tau_n\}$  be the initial  $n$  sample and  $\tau_{(1)} < \dots < \tau_{(n)}$  its order statistics. There is an upper part of the sample based on a threshold value  $\tau_{thr}$ ,  $T_{idle} = \{\tau_{(n-i+1)}\}_{i=1, \dots, k}$ , where  $\tau_{(k)} < \tau_{thr}$  and  $\tau_{(k+1)} \geq \tau_{thr}$ , that follows an

Exponential distribution,  $Exp(\lambda_{idle})$ . We divide therefore the trace into two parts, the first  $T_{idle}$  corresponds to the extreme events of the sample that follow an Exponential distribution and the second,  $T'_{idle}$ , comprises the rest of the sample that follows a mixture of  $K$  Exponential components under the MAP framework, whose parameters will be determined from the *EM* algorithm described in Section 2.4. The “idle” rate is estimated prior from the other parameters of the MAP model, and then is embedded to the parameter set so that the state-space of the MAP is extended by one,  $K^* = K + 1$ . We should recall that the sample  $T$  is drawn from a Hyper-exponential distribution under the MAP model, and we believe that the adoption of a mixture of Exponential distributions for the approximation of a sample that includes extreme events that characterize the tail of the distribution is not deemed unreasonable (Feldmann and Whitt, 1998).

So, we fit an Exponential distribution to inter-event times larger than a threshold value,  $\tau_{thr}$ , and the fitted parameter  $\hat{\lambda}_{idle}$  is the maximum likelihood estimator of the distribution. The threshold does not need to be the smallest value for the assumption to hold, it can be chosen quite conservatively aiming to fit the few observations at the tail of the inter-event times distribution but still to include enough data to allow a robust statistical analysis. Finch et al. (1989) have showed that for fitting mixture distributions the determination of the mixture proportion is crucial to converge to the global maximum in iterative algorithms where a starting vector of points is needed. Therefore, we implemented the following procedure to determine the minimum threshold. Firstly, we need to define the minimum threshold,  $\tau_{thr}^{min}$ , which we find reasonable to set as the mean value of the trace  $Tr$  to exclude the bulk of data (inter-event times) that do not characterize the tail of the distribution (extreme events). Then, we create an increasing set of  $\tau_{thr}$  values

$$\mathbf{T}_{thr} = \{\tau_{thr}^{min}, \tau_{thr}^{min} + s, \dots, \tau_{thr}^{max}\},$$



## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

with an assigned step,  $s$ , chosen arbitrarily, until no less than 30 events exist. Next, we move to the second part of the estimation procedure, where we compute the LL value of the MAP for each adopted  $\tau_{thr} \in \mathbf{T}_{thr}$  value. We know that in general when we use the *EM* algorithm the convergent LL value is highly dependent on the initial selection of the parameter set, especially as the number of latent states is getting higher (more local maximum points). Since there is no standard method to define the initial set, we followed a systematic procedure which is based on the k-means clustering algorithm, following a similar procedure with [Okamura and Dohi \(2009\)](#). The algorithm groups the trace into  $K$  clusters,  $\{\mathbf{c}_1, \dots, \mathbf{c}_K\}$ , where  $K$  corresponds to the number of MAP states and then, the diagonal elements of rate matrix  $\mathbf{D}_0$ ,  $\lambda_1, \dots, \lambda_K$ , are estimated from the inter-event times of each cluster assuming  $K$  Poissonian distributions,  $P(\lambda_i)$ ,  $i = 1, \dots, K$ . So, an initial starting vector for the implementation of the *EM* algorithm concerning the rate matrix,  $\mathbf{D}_0$ , is the following

$$\{\lambda_1^0, \dots, \lambda_K^0\} = \left\{ \frac{1}{\mathbf{c}_1}, \dots, \frac{1}{\mathbf{c}_K} \right\}.$$

For each repetition of the algorithm, new diagonal elements are drawn uniformly from the 99% confidence bounds of the Poisson distribution

$$\lambda_i^{new} = \lambda_i^0 \pm z_{a/2} \sqrt{\frac{\lambda_i^0}{n_i}},$$

where  $n_i$  is the sample length of cluster  $\mathbf{c}_i$ , and  $z_{a/2}$  corresponds to the standard normal distribution for  $a = 0.01$ . The rest of the parameters, i.e., elements of rate matrix,  $\mathbf{D}_1$ , are derived from a random transition probability matrix,  $\mathbf{P}$ , of the embedded Markov Chain of the model. The procedure is repeated many times ( $N_0$ ) and the  $N_{best}$  maximum LL values along with their corresponding parameter sets are stored and used as input to the *EM* algorithm (Algorithm 1). The BuTools program package ([Bodrog et al., 2014](#)) is used in the MATLAB environment for the implementation of the *EM* algorithm. Essentially, a



## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

grid-based procedure is followed around the estimated values of the occurrence rates. Finally, we choose as optimal  $\tau_{thr}$  point the one that corresponds to the maximum LL value. At the E-step of the *EM* algorithm the log-likelihood function is evaluated at the observation vector,  $Tr$ , but the estimated value of the idle state,  $\lambda_{idle}$ , at the diagonal matrix,  $\mathbf{D}_0$ , is considered a constant that we do not update at the M-step where the maximum likelihood estimates are computed.

---

### Algorithm 1 Estimation of “idle” occurrence rate and EM implementation

---

- 1: **Input:**  $Tr = \{\tau_1, \dots, \tau_N\}$  trace
  - 2: **Output:**  $\text{MAP}(\mathbf{D}_0, \mathbf{D}_1)$
  - 3: Set initial  $\tau_{thr}$ :  $T_{idle} = \{\tau_i : \tau_i > \tau_{thr}\} = \{\tau_i^{new}, \dots, \tau_{N_{idle}}^{new}\}$
  - 4: Fit an Exponential distribution to  $T_{idle} \rightarrow \hat{\lambda}_{idle} = \frac{1}{\hat{\tau}^*}$ , where  $\hat{\tau}^*$  mean value of trace  $T_{idle}$
  - 5: K-means algorithm to  $T'_{idle} = \{\tau_i : \tau_i \leq \tau_{thr}\}$  inter-event times that correspond to the  $K$  “no idle” states  $\rightarrow$  clusters  $\mathbf{c}_1, n_1, \dots, \mathbf{c}_K, n_K$ , equal to the number of states
  - 6: Compute the occurrence rates,  $\lambda_i^0 = \frac{1}{c_i}$ ,  $i = 1, \dots, K$
  - 7: **Repeat:**
  - 8: Update  $\lambda_i^{new} = \lambda_i^0 \pm U(-c^*, c^*)$ , where  $c^* = z_{\alpha/2} \sqrt{\frac{\lambda_i^0}{n_i}}$  the upper constraint
  - 9: Generate a random probability transition matrix  $\mathbf{P} \rightarrow \mathbf{D}_1 = -\mathbf{D}_0 \cdot \mathbf{P}$  under conditions
  - 10: i) irreducible and ii) aperiodic
  - 11: Compute  $LL_{new} = \log(L(\{\hat{\boldsymbol{\theta}}, \hat{\lambda}_{idle}\}/T))$
  - 12: **if**  $LL_{new} > \min_{i=1, \dots, N_{best}} LL_i$  **then**
  - 13:     store parameter set  $\hat{\boldsymbol{\theta}}_{new}$  and  $LL_{new}$
  - 14: **else**
  - 15:     reject parameter set
  - 16: **end if**
  - 17: **Until:** Maximum number of iterations,  $N_0$ , is reached
  - 18: Implement EM algorithm for the  $N_{best}$  parameter sets
  - 19: **Return:** Optimal parameter set  $\{\boldsymbol{\theta}, \lambda_{idle}\}$  after the EM algorithm convergence
- 

For instance, let's assume a process with  $K = 2$  states and rate matrices

$$\mathbf{D}_0 = \begin{bmatrix} -\lambda_1 & 0 \\ 0 & -\lambda_2 \end{bmatrix} \text{ and } \mathbf{D}_1 = \begin{bmatrix} q_{11}(1) & q_{12}(1) \\ q_{21}(1) & q_{22}(2) \end{bmatrix}.$$

Then, we add an “idle” state with total occurrence rate,  $\lambda_{idle}$ , so the new MAP will have  $K^* = 3$  states with new transition rate matrices

$$\mathbf{D}_0 = \begin{bmatrix} -\lambda_{idle} & 0 & 0 \\ 0 & -\lambda_1 & 0 \\ 0 & 0 & -\lambda_2 \end{bmatrix} \text{ and } \mathbf{D}_1 = \begin{bmatrix} q_{1*1}(1) & q_{1*2}(1) & q_{1*3}(1) \\ q_{21*}(1) & q_{22}(1) & q_{23}(1) \\ q_{31*}(1) & q_{32}(1) & q_{33}(1) \end{bmatrix}.$$

### 4.3.2 Pseudo-prospective forecasting framework

The two main statistical inference approaches for forecasting are the alarm and the probabilistic based forecasts. The former relies on a binary output, i.e., an alarm is switched on when the value of an alarm function exceeds a threshold and vice versa (Lippiello et al., 2012; Zechar and Jordan, 2008). The probabilistic forecasts consider a distribution for the earthquake frequency in the space-time-magnitude domain yielding occurrence probabilities, and they can be divided into grid-based and catalog-based forecasts. The methodology of the Collaboratory for the Study of Earthquake Predictability uses grid-based probabilistic forecasts implementing likelihood-based tests (Rhoades et al., 2011; Schorlemmer et al., 2007) for the comparison among the testing models. The main assumption is that the  $n$  space-time-magnitude testing bins of the experiment are considered independent, and earthquakes follow a Poisson distribution with parameter  $\lambda_i$  at each bin  $i$  based on the testing model  $\Lambda$ . Finally, a log-likelihood score is obtained by summing over the testing bins,  $LL = \sum_{i=1}^n \ln \frac{e^{-\lambda_i} \lambda_i^{x_i}}{x_i!}$ , with  $\mathbf{X} = \{x_i | i = 1, \dots, n\}$  the observed catalog, which is subsequently compared with different models.

However, the Poisson assumption for the earthquake distribution inside the testing bin might be unrealistic in cases where dependencies among events exist and additionally can lead to false rejection of models with overdispersion characteristics since the Poisson model considers its mean and variance values to be equal. Nandan et al. (2019) have shown that catalog-based probabilistic forecasts that consider the full distribution of earthquake number provide better

## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

results than the forecasts that are forced to use only the mean estimate rate under the Poisson assumption for each forecasting bin. In catalog-based forecasts simulated catalogs are provided for each bin under the testing model and the probabilities,  $p(Y_i)$ , that  $Y_i$  earthquakes occur inside the testing bin  $i$ , are used to obtain the log-likelihood score  $LL = \sum_{i=1}^n \log(P(Y_i))$ . This type of forecast assumes that simulations of the testing model should be converging to the observations if the model is the one that generates the data. Recently, the CSEP has developed new evaluation tools for forecasts specified as catalog-based and applied them to UCERF3-ETAS during the 2019 Ridgecrest sequence [Savran et al. \(2020\)](#). In the sequel, we will apply the catalog-based forecasting for the evaluation of the testing model, the Markovian Arrival Process.

One way to assess probability forecasts is based on the idea of defining a likelihood score in terms of consistency among forecasted and observed number of earthquake rates during a testing period  $T$ . The testing period is divided into  $n$  forecasting bins, such that  $T_i = [t_{i-1}^*, t_i^*)$ , with  $t_i^* = t_{i-1}^* + \delta$  for  $i = 1, \dots, n$ , and  $t_0^* = T_0$ , where  $T_0$  is the ending time of the learning period and  $\delta$  the length of the testing interval. If we denote as  $Y_i$  the variable that indicates the number of forecasted events at the  $i$ -th interval,  $T_i$ , then we can assign the forecasting probabilities  $p_{i,k} = P(Y_i = k)$  with  $\sum_{k=1}^{n_{max}} p_{i,k} = 1$  for all the testing bins, where  $n_{max}$  is an upper threshold for the maximum number of events. To compute the probabilities  $p_{i,k}$ , the model can use information only from the training period. In our procedure, we start with the first training period until time  $T_0$ , which is then updated by  $\delta$  time units until  $T_0 + \delta$ , to compute the forecasting probabilities of the testing interval  $[T_0 + \delta, T_0 + 2\delta)$  and so on. If the observed number of events is equal to  $X_i$ , then we can define the forecasting score as

$$S_i = \sum_{k=0}^{n_{max}} \log(p_{i,k}) I_{X_i}(k), \quad (4.1)$$

where  $I_{X_i}(k) = \begin{cases} 0 & X_i \neq k \\ 1 & X_i = k \end{cases}$  for the interval  $T_i$ , which is the logarithmic probability to forecast the observed number of events and its negative value is known as entropy score (Daley and Vere-Jones, 2004).

It is easy to expand the definition to the total testing period,  $S = \sum_{i=1}^n S_i = \sum_{i=1}^n \sum_{k=0}^{n_{max}} \log(p_{i,k}) I_{X_i}(k)$ . It represents the logarithmic probability to forecast the observed number of events at each interval  $T_i$ , i.e.,  $\log(P(Y_1 = X_1, \dots, Y_n = X_n))$ . In the case, where the forecasting probabilities are the real ones of the physical process, the log-likelihood score gets its maximum value which is equal to zero. We are interested in the information that we gain against a reference model, with corresponding forecasting probabilities  $p_{i,k}^{ref}$ . Thus, we define the quantity,

$$D = S - S_{ref} = \sum_{i=1}^n \sum_{k=0}^n \log(p_{i,k}/p_{i,k}^{ref}), \quad (4.2)$$

which is essentially the probability gain of the tested model against the reference one.

Next, we need to estimate the forecasting probabilities,  $p_{i,k}$ , for each interval  $T_i$ . We proceed with  $K$  simulations of the corresponding model  $\mathbf{\Lambda}$  up to the end of each testing interval  $T_i$ . Each simulation produces one possible scenario for each forecasting interval  $T_i = [t_{i-1}^*, t_i^*)$ , so after the implementation of  $K$  simulations we define as  $N_i$  the number of realizations with  $X_i$  events (successes), which is the observed number of events, within the interval  $T_i$ . The forecasting probability estimates are then given by the fraction over all the simulations and are defined as  $\hat{p}_{i,k} = \frac{N_i}{K}$ ,  $i = 0, \dots, n$ . Essentially, we calibrate forecasting probabilities over multiple testing intervals based on simulations generated from the tested model and we assume that they should converge to the observations if the model is the "true" one.

### 4.3.2.1. Competing forecasting models

Figure 4.4 shows an example of a data set that includes the training period up to time  $T_0$  and the first testing bin  $T_1$ .

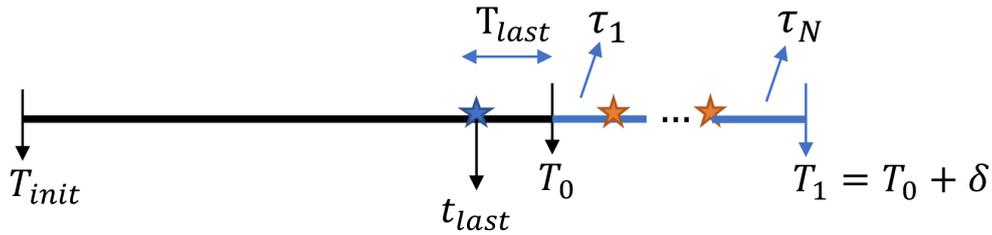


Figure 4.4: An example that includes the training period  $[T_{init}, T_0)$ , the occurrence time of the last event in the training period,  $t_{last}$ , with the blue star and the first testing bin  $T_1 = [T_0, T_0 + \delta)$  along with the observed seismicity (yellow stars).

For the generation of the  $K$  simulations with the MAP, we proceed as follows. Firstly, the model parameters are calibrated based on the observations up to time  $T_0$ , which coincides with the starting time of the testing period. Then, the state probabilities of the hidden process  $J_t$ ,  $p_i(T_0) = P(J_{T_0} = i)$ , are estimated based on the forward and backward equations through Equation 2.14 which we consider as the initial probability vector for the simulation procedure. This is equivalent with the state probability at the last earthquake occurrence of the learning period, at  $t_{last}$  in Figure 4.4, since transitions among the hidden states coincide with the occurrence of earthquakes, namely  $p_i(t_k) = p_i(t)$  for  $t_k \leq t < t_{k+1}$ . Then, an Exponential distribution is assumed for the generation of the time until the first event,  $\tau_1$ , with parameter that depends on the estimated state of the hidden process at time  $t_{last}$ . Details on the simulation from the MAP model are given in Section 2.5. Here, we sum up the simulation procedure for forecasting purposes in the following steps:

- Set  $J(T_0) = \operatorname{argmax}_{1 \leq i \leq K} p_i(T_0)$  where  $\mathbf{p}(T_0)$  the initial probability vector.



## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

---

- Let us denote  $J(T_0) = i$  the most probable state. The sojourn time,  $\tau$ , until the next earthquake occurrence is extracted by an Exponential distribution with parameter  $\{-D_0\}_{ii} = \lambda_i$  through Equation (2.12).
- Generate the state of the hidden process at time  $T_0 + \tau$ ,  $J(T_0 + \tau)$ , based on the conditional distribution  $P(J(T_0 + \tau)/J(T_0))$ , given by Equation (2.2), i.e.,  $\{\frac{q_{i1}(1)}{\lambda_i}, \dots, \frac{q_{iK}(1)}{\lambda_i}\}$ .
- Update time  $t = T_0$  by  $t = T_0 + \tau$  and go to step 2.
- Continue until the end of the testing period  $T_1$ .

For the Poisson model with parameter  $\lambda_{pois}$ , the inter-event times follow an Exponential distribution. Therefore, the number of events in the testing bin  $T_1$  is independent of the time elapsed since the last event,  $T_{last}$ , in the training period, i.e.,  $P(T \leq \tau + T_{last}/T > T_{last}) = P(T \leq \tau)$ . The simulation of the earthquake occurrence times is a typical procedure based on the thinning method (Equation 2.12).

Concerning Gamma and Weibull renewal models, the time to the next earthquake depends on the elapsed time since the last event,  $T_{last}$ . Therefore, for the generation of the first event into the testing interval  $T_1$  we should consider the truncated distribution of the Gamma and Weibull models, respectively. In particular, the truncated Weibull cumulative distribution from the left, with  $a$  the shape and  $\beta$  the scale parameters, respectively, is easily derived using the Bayes theorem and has the following form

$$P(x \leq T_{last} + T/x > T_{last}) = 1 - e^{-\frac{T_{last}^a - (T_{last} + T)^a}{\beta^a}}, \quad T \geq 0.$$

The truncated Gamma distribution,  $G(a, b)$ , has a more complex functional form with corresponding density

$$f(x) = \frac{b^a}{\Gamma(a, bT_{last})} e^{-bx} x^{a-1}, \quad x \geq T_{last}.$$



## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

---

For the subsequent events inside the testing bin the corresponding distribution functions of the models are used. Finally, we generate forecasts under the temporal ETAS model, which is considered the standard model for aftershock forecasting. The main assumption of the model is that there is a constant loading of independent events under a Poisson process with rate  $\mu$  and that each earthquake generates its own aftershocks under the triggering term  $\lambda_j(t)$  given by Equation A.1. The ETAS model is proven appropriate for aftershocks modeling, whereas in our study it is fitted to data sets of large earthquakes where its empirical laws incorporated into the triggering term (the Omori law for the aftershocks temporal decay and the productivity law for the expected number of direct aftershocks) might not be appropriate. The procedure to produce forecasts is similar to the one implemented at [Nandan et al. \(2019\)](#).

In step 1 the earthquakes of the training set are considered mother events that can create daughters inside the testing period. Consequently, these daughters can trigger their own ones and so on. To obtain the magnitudes and the times of these earthquakes the simulation procedure given in Appendix A.3 is followed. In particular, the first-generation of aftershocks for all the earthquakes in the training period are generated. Each earthquake with magnitude  $m_i$  triggers aftershocks according to a Poisson process with rate  $k(m_i)$  (Equation A.2). The times of those aftershocks are generated from the Omori law distribution (Equation A.3). The magnitudes are independent from the events temporal distribution and follow the GR law truncated from the left at the completeness magnitude,  $m_{thr}$ . The functional form of their distribution is given by Equation A.7, and the  $b$ -value of the GR law is estimated from the training data set. For next generation aftershocks, the triggering step is repeated until there are no more generated events. Events that exceed the time of the testing period are neglected.

In step 2 we generate the background earthquakes that are expected to occur during the testing period as well as the cascade of aftershocks that are triggered by these background events. We produce independent events according

to a Poisson distribution with mean value equal to the estimated background rate,  $\mu$ , of the ETAS model. Therefore, inter-event times are simulated from an Exponential distribution  $Exp(1/\mu)$ , until they exceed the ending time of the testing period,  $t_{end}$ . We then simulate the triggered aftershocks corresponding to all the background events as described at step 1 and store them along with the ones produced during step 1.

## 4.4 Application in circum-Pacific belt

In this section we show that the temporal behavior of the large earthquakes in circum-Pacific belt cannot be captured well by the MAP model due to the presence of short inter-event times and the brevity of the earthquake catalog. Subsequently, we proceed with the application of the MAP model with the idle state. We investigate the non-stationary features of the data sets with  $M \geq 7.0$ , 7.6, 7.7, 7.8, and we perform pseudo-prospective experiments to evaluate its performance in forecasting the full distribution of earthquakes number. We compare its efficiency to renewal models and the temporal ETAS model.

### 4.4.1 Application of MAP

For the data sets with  $M \geq 7.6$ , 7.7, 7.8, MAP models with 2 and 3 states are fitted to the data, whereas for the data set with  $M \geq 7.0$  an additional MAP with 4 states is considered. The temporal evolution of seismicity is more complex for the last data set as short-term clustering is more intense, so more states might be needed. We estimate the parameters through the maximization of the log-likelihood function given in Equation 2.6 with the use of the *EM* algorithm and a grid-based procedure for the determination of the algorithm's initial input described in Section 4.3.1. In all cases a MAP model with 3 states seems sufficient for the description of the earthquake temporal distribution

CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING  
LARGE EARTHQUAKES NUMBER

according to the AIC values. In Appendix C.2 we give the supplementary results concerning the LL and AIC values of the fitted models as well as the goodness of fit results.

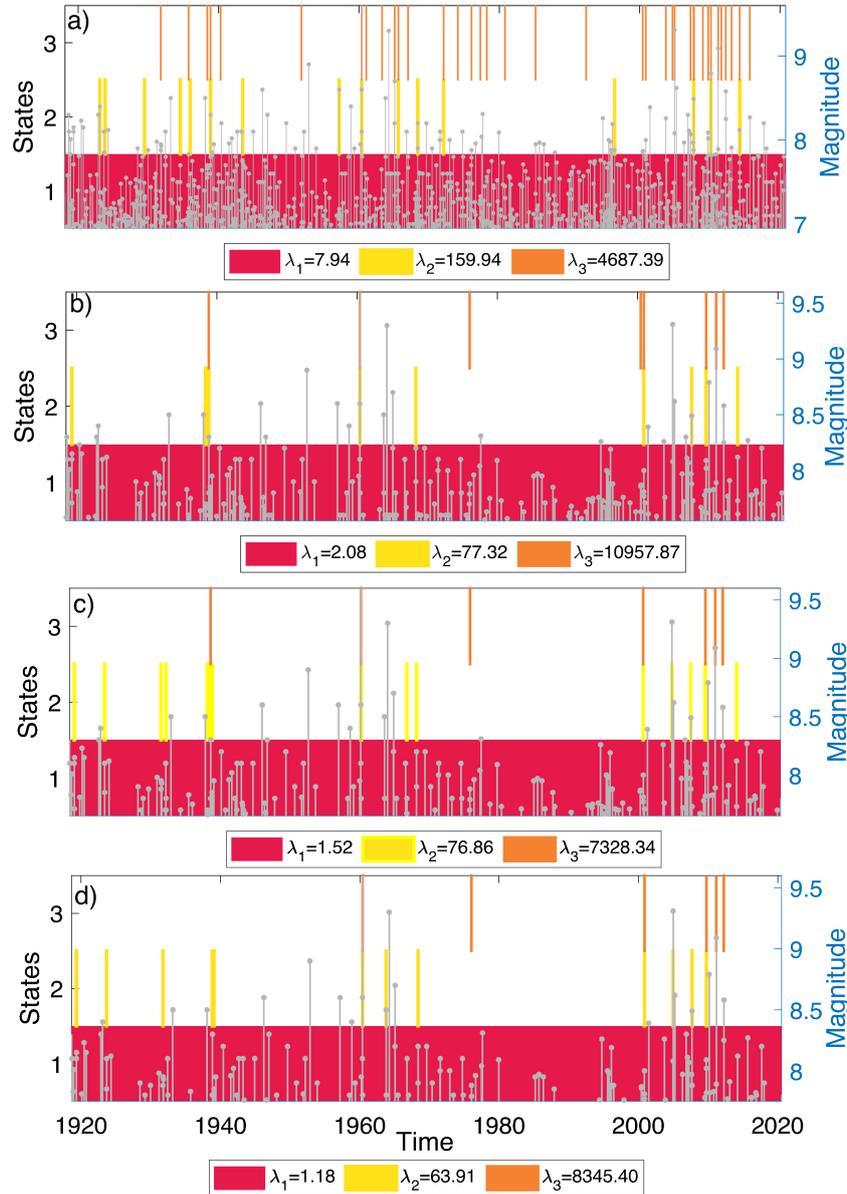


Figure 4.5: The most probable path of the hidden states along with the earthquake magnitudes (grey vertical lines) as a function of time for the data sets with  $M \geq 7.0$  (a),  $M \geq 7.6$  (b),  $M \geq 7.7$  (c) and  $M \geq 7.8$  (d). Red color corresponds to state 1, yellow color to state 2 and orange color to state 3 for all the panels, respectively.

Figure 4.5 shows the most probable sequence of transitions among the

latent states of the MAP model for the four data sets based on the evaluation of the state probabilities,  $p_i(t)$ ,  $i = 1, \dots, K$ , (Equation 2.14). In all cases, the estimated hidden path is dominated by the lowest seismicity rate,  $\lambda_1 = 7.94, 2.08, 1.52, 1.18$  events/yr (red color), for  $M_{thr} = 7.0, 7.6, 7.7, 7.8$ , respectively. This is either an indicator of stationarity for the background seismicity or the inadequacy of the model to capture the long-term seismicity changes. The estimated rates of state 1 are higher than the empirical occurrence rates in the upper part of the inter-event times distribution,  $\hat{\lambda} = 3.38, 0.9, 0.66, 0.51$ , i.e., the events corresponding to the fourth quantile (Figure 4.3). Finally, the other two states seem to correspond to the apparent short-term clustering.

To verify the influence of the short-term clustering to the fitting results of the MAP model, we proceeded to the separation of the clustered from the background seismicity based on the implementation of the Nearest-Neighbor algorithm (Zaliapin and Ben-Zion, 2013a) to the data sets with  $M_{thr} = 7.0$  and  $M_{thr} = 7.6$ . This is a clustering algorithm that is based on the space-time-magnitude distance metric among two earthquakes given by Equation B.4 by Baiesi and Paczuski (2004). The details of the method are given in Appendix B.3. There are only two free parameters, the spatial fractal dimension  $d_f$  and the  $b$  value, which are considered equal to  $b = 1.05, 1.28$  for earthquakes with  $M \geq 7.0$  and  $M \geq 7.6$ , respectively, and  $d_f = 1.29$ , for both thresholds. The logarithm of the separation distance is equal to  $\log \eta_0 = -5.80$  (Figure 4.6a) and  $\log \eta_0 = -7.67$  (Figure 4.6c), respectively, based on the intersection of the two modes in the 1D density distribution of distances.

The resulted declustered earthquake catalogs comprise 697 events with  $M \geq 7.0$  and 195 events with  $M \geq 7.6$ , respectively. Then, MAP models with 2 and 3 states are fitted to the data, and the two-state MAP models are considered optimal based on the AIC values. Figure 4.7a,c illustrate the transitions among the two states of the MAP models for the data sets with  $M \geq 7.0$  and  $M \geq 7.6$ , respectively. After the removal of the short-term clustering ef-

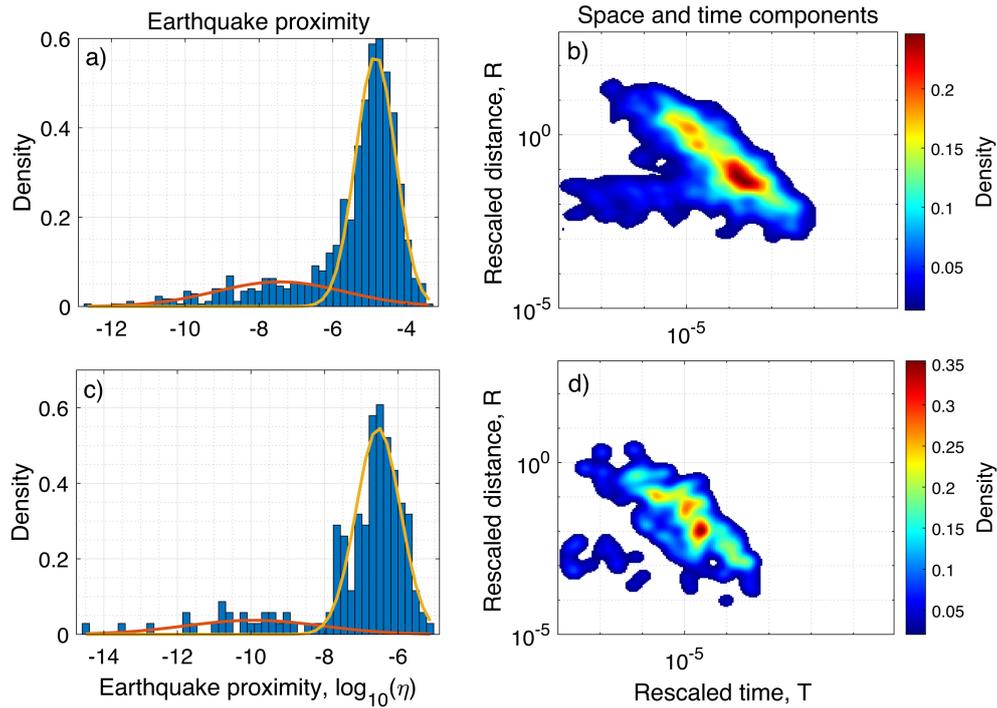


Figure 4.6: 1D density distribution of  $\log \eta$ , with estimated Gaussian densities for clustered (yellow) and background (orange) components for the earthquakes with  $M \geq 7.0$  a) and  $M \geq 7.6$  c), respectively. 2D joint distribution of rescaled space and time distances among all pair of events for the earthquakes with  $M \geq 7.0$  b) and  $M \geq 7.6$  d), respectively.

fect with the NN algorithm, two main observations can be made. Firstly, the background seismicity seems to be characterized by non-stationarity as the two states are alternating for both data sets in contrast with the initial data sets where a single state (state 1) is associated to the background seismicity (red color in Figure 4.5). The intense short periods with the high seismicity rates (yellow and orange colors in Figure 4.5) are replaced by a state with a low seismicity rate,  $\lambda_1 = 5.06, 1.58$ , for  $M \geq 7.0$  and  $M \geq 7.6$ , respectively (red color in Figure 4.7a,c) that is closer to the empirical rates of the long inter-event times distribution. This can also be seen from the estimated intensity function of the two models,  $\lambda(t)$  (Equation 2.15), in Figures 4.7b,d. Long-term variations are observed, with three common periods of seismic quiescence for the two data sets, during 1948-1963, 1978-1992 and 2002-2004, respectively. So, the constant background rate is not a feature of the data. Secondly, we can

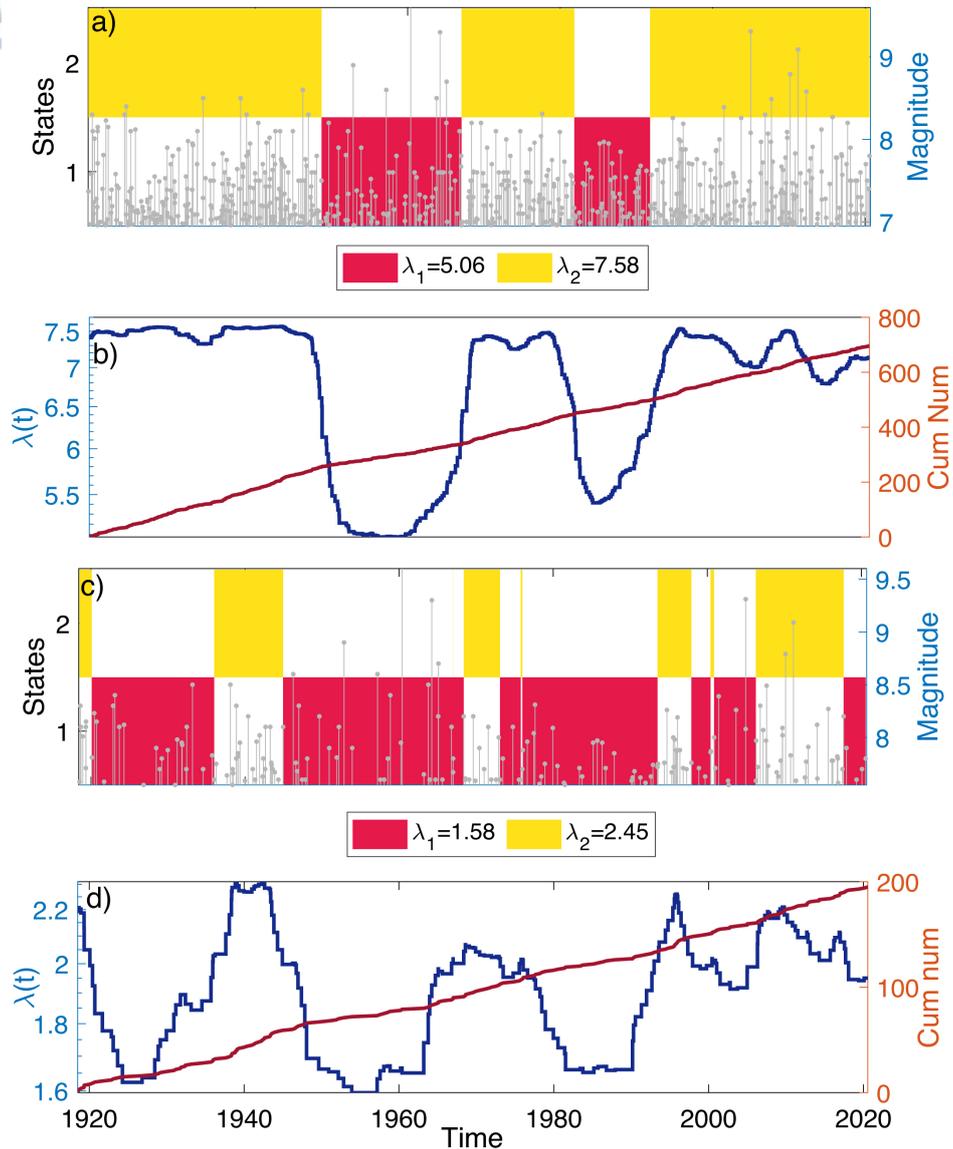


Figure 4.7: The most probable path of the hidden states along with the earthquake magnitudes as a function of time for the data sets with  $M \geq 7.0$  (a) and  $M \geq 7.6$  (c). Red color corresponds to state 1 and yellow color to state 2. The expected seismicity rate is denoted by the blue continuous line and the cumulative number of events with the maroon line, as a function of time for the data sets with  $M \geq 7.0$  (b) and  $M \geq 7.6$  (d).

conclude that the MAP models fitted to the initial earthquake catalogs cannot capture the long-term changes in seismicity due to the existence of short-term events.

### 4.4.2. Application of MAP with idle state

#### 4.4.2.1 Evaluation

Firstly, we investigated whether a stationary Poisson model can adequately describe the earthquake occurrence patterns in each data set. We know that statistical tests cannot always reveal the non-stationarity of large earthquake catalogs due to data shortage (Daub et al., 2015). We believe, however, that it is important to investigate the goodness of fit to the data of the simplest model before proceeding to more refined models. We tested whether the inter-event times follow the Exponential distribution with the Lilliefors goodness of fit test (Lilliefors, 1969), which is an analog of the KS test when the parameters of the null hypothesis distribution,  $F_0(x)$ , need to be estimated from the sample and the  $p$  values are computed through Monte Carlo simulations (Section 2.8). Next, the Runs test is applied to decide if correlations exist among them and the resulted  $p$  values are given in Table 4.1. We note that this is a common procedure to detect non Poissonian behavior (Lombardi and Marzocchi, 2007; Touati et al., 2016). Concerning, the data set with  $M \geq 7.6$  even though, we cannot reject the hypothesis of randomness ( $p = 0.08$ ), its value is relatively low providing evidence against the null hypothesis. The Exponential distribution does not adequately fit the observations ( $p = 0.003$ ). Similar results are obtained for earthquakes with  $M \geq 7.0$ ,  $M \geq 7.7$  and  $M \geq 7.8$ . A common feature for all data sets is that the Poissonian behavior is rejected with very low  $p$ -values (Table 4.1). This implies that a renewal process with non-exponential times might be more appropriate than a stationary Poisson process for the approximation of the temporal distribution of the earthquakes. Therefore, Weibull and Gamma distributions which are widely used to study quasi-periodic earthquake occurrences are also tested as null hypothesis. Again, the  $p$  values ( $p = 1E - 5, 2E - 5$ ) for  $M \geq 7.6$  suggest that the null hypothesis cannot be accepted for both distributions. Very low values are also derived for the other 3 data sets.

CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING  
LARGE EARTHQUAKES NUMBER

Table 4.1: Statistics and MAP fitting evaluation for the earthquakes with  $M \geq 7.0$ ,  $M \geq 7.6$ ,  $M \geq 7.7$  and  $M \geq 7.8$ . The  $p$  values of the Runs test for the inter-event times are given in column 4 and the  $p$  values of the Lilliefors test under the null hypothesis of the Exponential, Weibull and Gamma distributions are given in columns five, seven and nine, respectively.

$M_{thr}$	$N_0$	Rate $yr^{-1}$	Runs Test	Exponential		Weibull		Gamma	
				$Lill_{test}$	$\Delta AIC$	$Lill_{test}$	$\Delta AIC$	$Lill_{test}$	$\Delta AIC$
7.0	909	8.84	0.80	0	145.6	1E-4	29.0	5.5E-3	0
7.6	239	2.32	0.08	3.1E-3	50.4	1E-5	17.7	2E-5	2.4
7.7	178	1.73	0.36	2.0E-4	49.83	1.4E-3	11.77	1.9E-2	0
7.8	138	1.34	0.86	3.8E-3	44.46	4.2E-3	9.85	1.9E-2	0
$MAP_3$									
	$\lambda_{idle}$	$\lambda_1$	$\lambda_2$	$\Delta AIC$	Runs Test	KS			
7.0	2.90	8.54	1889	4.9	0.39	0.12			
7.6	1.13	2.59	1080	0	1.00	0.76			
7.7	0.78	1.93	257	15.91	0.48	0.30			
7.8	0.61	1.62	818	14.07	0.90	0.36			

Moving to the MAPs fitting, we considered different values for the free parameter of Algorithm 1,  $\tau_{thr} = \{\bar{\tau}, \dots, \tau_{thr}^{max}\}$ , with corresponding  $\lambda_{idle} = \lambda(\tau_{thr})$  values. The optimal threshold corresponds to the maximum log-likelihood following the procedure described in Section 4.3.1. MAPs of two and three states are fitted to the data and the three states model with  $S = 0, 1, 2$ , is selected based on the minimum value of the Akaike Information Criterion (Equation 2.16). State 0 will be expressed as “idle” state henceforth.

According to the occurrence rates ( $\lambda_{idle}, \lambda_1, \lambda_2$ ) shown in Table 4.1, idle state implies the existence of relative seismic quiescence periods with seismicity rate equal to  $\lambda_{idle} = 1.13 (yr^{-1})$  and corresponding expected sojourn time equal to 6.08 years for  $M \geq 7.6$ , whereas the middle state, state 1, indicates the presence of periods of relatively higher seismicity rates, with value  $\lambda_1 = 2.59 (yr^{-1})$  and sojourn time 4.42 years, again for  $M \geq 7.6$ . The state 2 might be related to regional triggering due to the very short sojourn periods (12.3 hours) of intense seismicity rate with value equal to  $\lambda_2 = 1080 (yr^{-1})$  for  $M \geq 7.6$ . The occurrence rates given in Table 4.1 (non-negative diagonal elements of rate matrix  $\mathbf{D}_0$ ), act as an index for the seismicity evolution for each data set, however, the elements of matrix  $\mathbf{D}_1$  determine the sequence of transitions

among the states. A thorough analysis based on the hidden paths will be given below.

We conduct a residual analysis for the MAP model fitting and apply the Runs and KS tests to the transformed successive inter-event times. The transformed times along with the 95% KS confidence bounds are presented in Figures 4.8a,b,e,f, along with the auto-correlation function which is computed and illustrated for a visual inspection in Figures 4.8c,d,g,h. The transformed times pass the randomness and goodness of fit tests with high  $p$ -values ( $p > 0.05$  shown in Table 4.1 for all data sets) and the residuals do not show significant discrepancies from the bisector that corresponds to the stationary unit Poisson process.

Next, we proceed to a quantitative comparison of the MAPs with the fitted renewal processes based on their LL values and their complexity (considering the number of their parameters). Table 4.1 shows for each model the difference between its AIC value from the minimum one,  $\Delta AIC = AIC - AIC_{min}$ . Zero value is assigned to the one with the minimum AIC value. We observe that the Gamma distribution returns similar or better values from the MAP model. For  $M_{thr} = 7.6$  the difference is equal to  $\Delta AIC = 2.4$  in favor of the MAP model (Table 4.1), whereas for the other data sets the differences are in favor of the Gamma and Weibull distributions. Gamma distribution suggests the existence of short-term clustering due to the factor  $x^{a-1}$  (higher probabilities for short inter-event times than the Poisson distribution) with an exponential decay for the long-term behavior of independent events. MAP is anticipated to be a more complex model due to the comparatively large number of parameters in relation with the limited data set. However, the embedded hidden states allow us to reveal additional details of the temporal patterns. Knowing the state probabilities,  $p_i(t)$ , and therefore the seismicity rate at time  $t$ , we can infer if there are systematic periods of seismic quiescence or short-term clustering and their expected duration. In Section 4.4.3 we show that this information might contribute to the forecasting skill of the MAP compared to the other models.

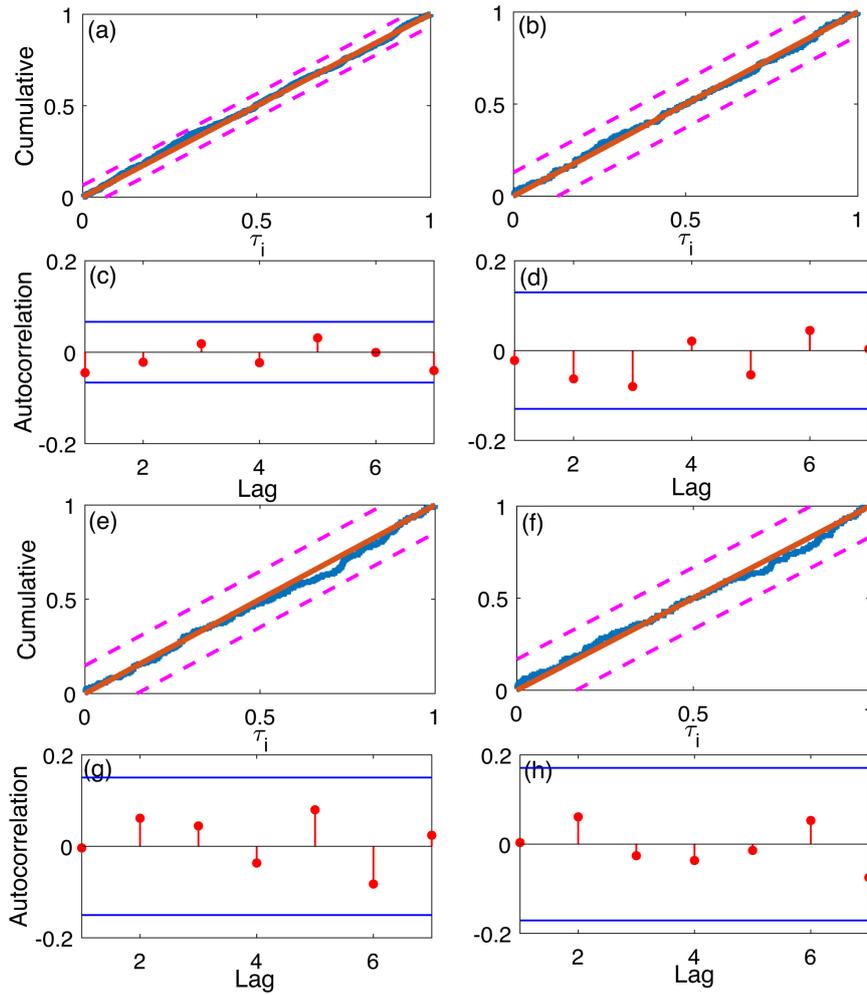


Figure 4.8: Comparison between the cumulative number of residuals (blue step function) and the stationary Poisson process with unit rate (brown line) (a) for threshold  $M_{thr} = 7.0$ , (b)  $M_{thr} = 7.6$ , (c)  $M_{thr} = 7.7$  and (d)  $M_{thr} = 7.8$ . The pink dashed lines indicate the 95% confidence bounds. Auto-correlation function of the  $\Delta\tau_i$  values (red vertical lines) and the blue horizontal lines indicate the corresponding confidence bounds for threshold (c)  $M_{thr} = 7.0$ , (d)  $M_{thr} = 7.6$ , (g)  $M_{thr} = 7.7$  and (h)  $M_{thr} = 7.8$ , respectively.

#### 4.4.2.2 Existence of non-stationarity

In what follows, the transitions among the three states with rates,  $(\lambda_{idle}, \lambda_1, \lambda_2)$  are evaluated through the state probabilities which are given by Equation 2.14. In addition, the expected seismicity rate at each time, namely the intensity function  $\lambda(t)$ , is estimated through Equation 2.15, and can be used as one more indicator for temporal changes in the seismicity. Figure 4.9 visualizes the evo-

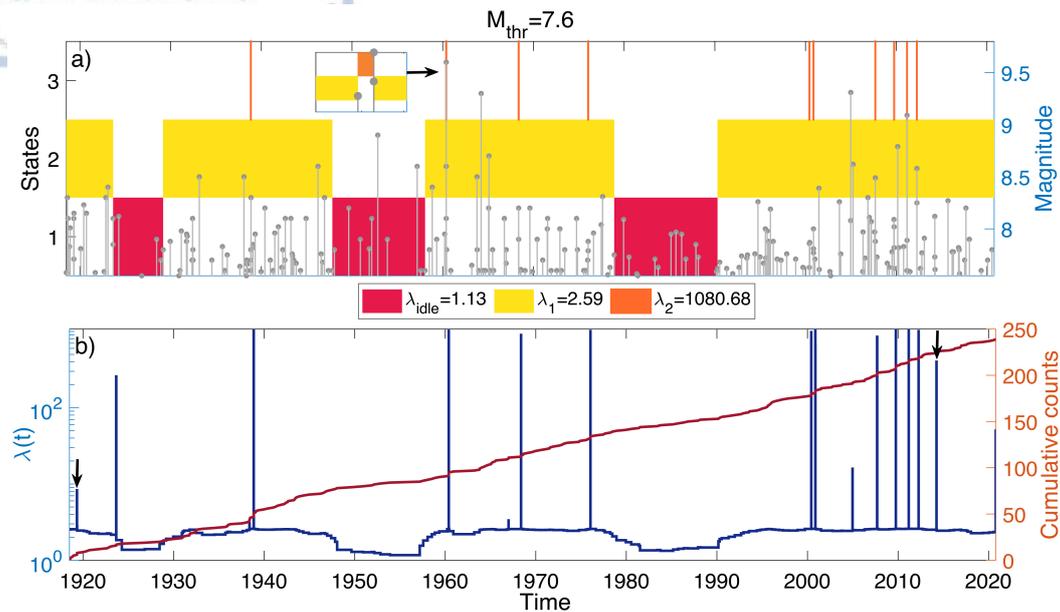


Figure 4.9: (a) The most probable path of the latent states along with the earthquake magnitudes of the events as a function of time with  $M \geq 7.6$ . Inset magnifies the transitions among states 1 and 2, which are otherwise difficult to visualize due to the short sojourn time in state 2. (b) The expected seismicity rate is denoted by the blue continuous line and the cumulative number of events with the maroon line, as a function of time. The highest peaks of the intensity rate correspond in most cases to events that occurred during the sojourn of the process in state 2. The black vertical arrows on panel (b) show two peaks of the intensity function that do not correspond to state 2, nevertheless, they are associated with regional earthquake interactions after the investigation of their spatial distribution. The first one corresponds to the 1919 Tonga earthquake with  $M_w$ 8.1 and the second one to the 2014 Iquique, Chile earthquake with  $M_w$ 8.1.

lution of the seismicity for the earthquakes with  $M \geq 7.6$ . In this way, we can explicitly depict the estimated seismicity rate during the entire period and reveal intervals of relative quiescence, as well as periods of excessive activity and of possible triggering.

For earthquakes with  $M_{thr} = 7.6$  we observe that the transitions between periods of relative seismic quiescence (red color in Figure 4.9a,  $\lambda_{idle} = 1.13$ ) and the ones of relatively high seismicity rate (yellow color Figure 4.9a,  $\lambda_1 = 2.59$ ) occur in a non-regular time scale varying from a couple of years to a couple of decades. Although there can be slight differences in the corresponding state

CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING  
LARGE EARTHQUAKES NUMBER

sequences of the higher magnitude thresholds ( $M \geq 7.7$ ,  $M \geq 7.8$  Figure 4.10), there are some persistent temporal intervals with increased seismicity rate (state 1, yellow color) such as between 1935-1947, 1958-1960, 1963-1978 and 1993-1997 (Figures 4.9 and 4.10). Common periods of relative seismic quiescence like during 1924-1928, 1953-1957 and 1980-1990 for all data sets are observed (Figure 4.9), increasing in number with the magnitude threshold (Figure 4.10), especially between 1977 and 2000, which is a period that coincides with an observed decreased moment rate (Figure 4.2b) mentioned also by [Zaliapin and Kreemer \(2017\)](#). The apparent increase in great ( $M \geq 8.0$ ) earthquakes occurrence since 2004 is also recognized as a period of relatively high seismicity in all cases (sojourn of the process in state 1-yellow color in Figures 4.9 and 4.10).

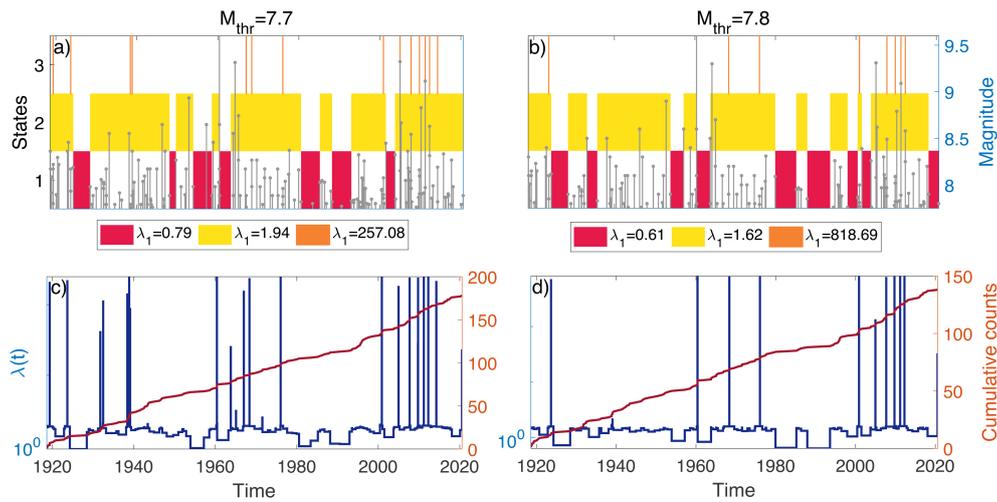


Figure 4.10: The most probable path of the latent states along with the magnitude distribution in time of the events with  $M \geq 7.7$  (a) and with  $M \geq 7.8$  (b). Red color corresponds to  $\lambda_{idle} = 0.79$ ,  $0.61$ , yellow color to  $\lambda_1 = 1.94$ ,  $1.62$  and orange color to  $\lambda_2 = 257.08$ ,  $818.69$  for the panel (a) and panel (b), respectively. (c), (d) The expected seismicity rate is denoted by the blue color and the cumulative moment of the events with the maroon.

Concerning the dataset with  $M_{thr} = 7.0$ , state 1 (yellow color in Figure 4.11,  $\lambda_1 = 8.54$ ) is dominant during the entire period, with short seismic quiescence periods during 1956-1957 and 1982-1985 (red color in Figure 4.11,  $\lambda_{idle} = 2.91$ )

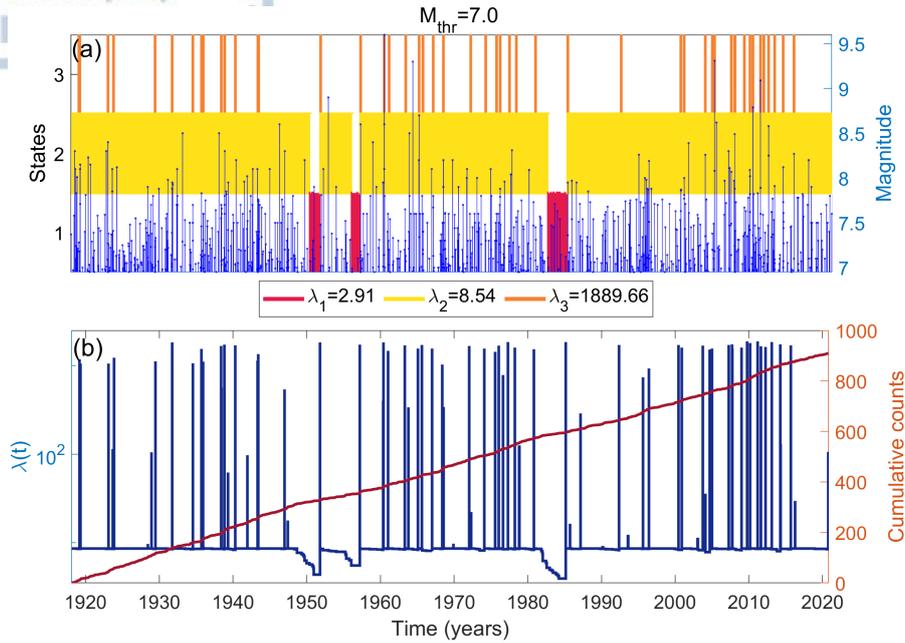


Figure 4.11: The most probable path of the latent states along with the magnitude distribution in time of the events with  $M \geq 7.0$ . Red color corresponds to  $\lambda_{idle} = 2.91$ , yellow color to  $\lambda_1 = 8.54$  and orange color to  $\lambda_2 = 1889.66$  for the panel (a). (b) The expected seismicity rate is denoted by the blue color and the cumulative number of the events with the maroon.

In Section 4.4.1 we show that after neglecting the clustered events from the initial dataset with  $M \geq 7.0$ , long-term changes in the remained seismicity are revealed by the MAP model (Figure 4.7), and two long seismic quiescence periods exist during 1948-1966 and 1981-1991 (red color in Figure 4.7a and blue step-function in Figure 4.7b). The abundance of smaller magnitude events compared to the dataset with  $M \geq 7.6$  increases the short-term clustering effect so that the MAP model cannot capture the long-term changes in the occurrence rate. Considering the existence of the idle state and implementing the two-step estimation procedure might not be appropriate for data sets with smaller magnitude cutoffs.

Clear evidence is provided from state 2 (yellow color in Figure 4.9a), which is the state with the highest corresponding Poisson rates. Events that occurred during the sojourn of the hidden process in state 2 seem to express the short-term localized clustering, which is visible in Figure 4.12 where we plot the

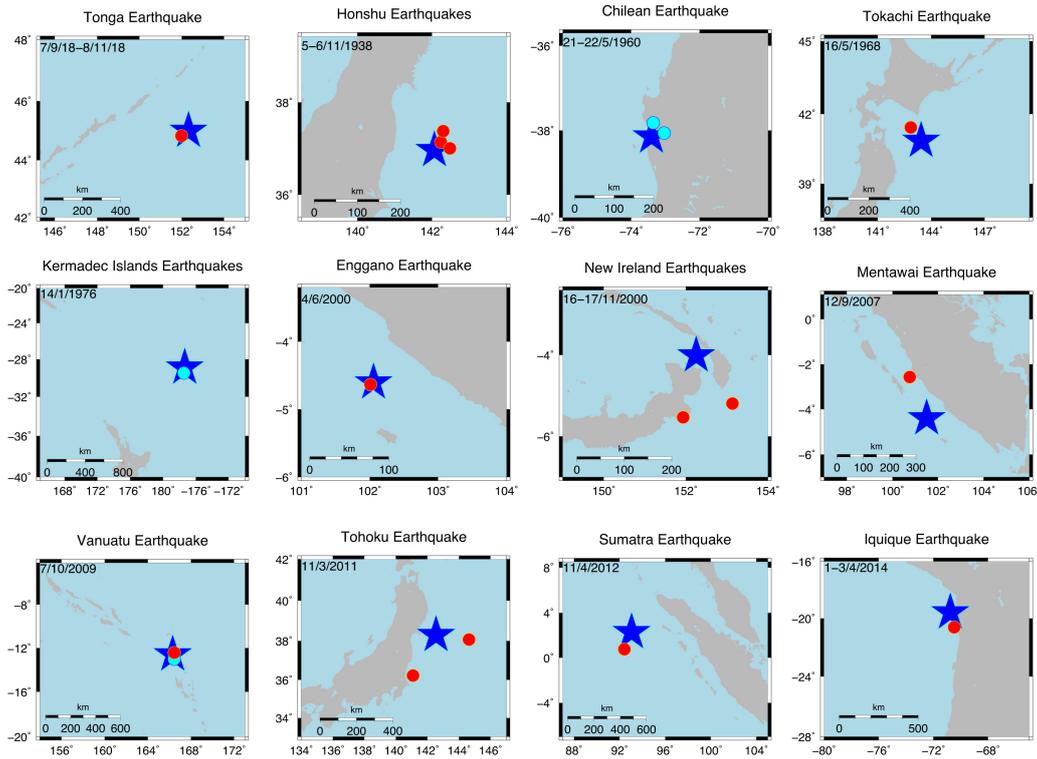


Figure 4.12: Epicentral distribution of the events with  $M \geq 7.6$  that occurred in the state with the highest seismicity rate,  $\lambda_2$  and two cases with high intensity values (black vertical arrows in Figure 4.9b). All cases correspond to regional spatiotemporal clustering. Blue stars correspond to the events with the largest magnitude, cyan and red circles to prior and subsequent events, respectively. The date at each panel represents the duration of each group of events.

epicenters of earthquakes with  $M \geq 7.6$  for each subset of consecutive events that occurred in this state. All cases correspond to large ( $M \geq 7.8$ ) main shocks (blue stars) and their triggered subsequent events (red circles), except the 1960 Chilean and 2009 Vanuatu earthquakes, where prior (cyan circles) events are also observed. In all eleven cases spatiotemporal clusters of known main shock-aftershock pairs are formed, evidencing that state 2 is related to the existence of spatiotemporal clustering. Considering the expected seismicity rate,  $\lambda(t)$ , (Figure 4.9b, blue step-function), we observe some peaks with  $\lambda(t) > \lambda_1$ , that correspond to regional earthquake interactions, however, without belonging to state 2. The two black vertical arrows in Figure 4.9b show

the 1919 Tonga earthquake (in Figure 4.12) with  $M_w 8.1$  and  $\hat{\lambda}_t = 8.28$  ( $yr^{-1}$ ) and the April 2014 Iquique, Chile earthquake (in Figure 4.12) with  $M_w 8.1$  and  $\hat{\lambda}_t = 399.99$  ( $yr^{-1}$ ).

### 4.4.3 Pseudo-prospective experiments

In this section we perform a comparison among the proposed MAP model with the idle state and the Exponential, Gamma and Weibull renewal models in terms of their forecasting skills. Additionally, we implement a sensitivity analysis on the forecasting interval and on different versions of the MAP model. We investigate whether the MAP model with the three embedded occurrence rates is efficient to forecast the large subduction earthquakes number and if it performs better than renewal processes with the Exponential, Gamma and Weibull distributed inter-event times. Firstly, the learning and testing periods are defined and then the MAP model along with the Poisson, Gamma and Weibull distributions are fitted to the observations of the learning period. We set the ending time of the learning period,  $T_0$ , equal to 31 December 2003, and the testing period from the beginning of 2004 until the end of 2020, lasting 17 years (Table 4.2).

Table 4.2: Details on the setup of the forecasting experiment for the data sets with  $M \geq 7.0$ ,  $M \geq 7.6$ ,  $M \geq 7.7$  and  $M \geq 7.8$ .

$M_{thr}$	Learning period [1918, $T_0$ ]	Testing period ( $T_0$ , 2020]	$T_0$	K
7.0	748	161		
7.6	190	49	31/12/2003	100000
7.7	135	35		
7.8	86	29		

Next, we proceed to the implementation of  $K = 100000$  simulations with each forecasting model and we evaluate their performance based on the score function defined in Equation 4.1. Figure 4.13 shows the score of the four models as a function of time obtained from 6-months, 1-year, 2-year, and 3-year long

CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING  
LARGE EARTHQUAKES NUMBER

experiments for earthquakes with  $M \geq 7.6$ .

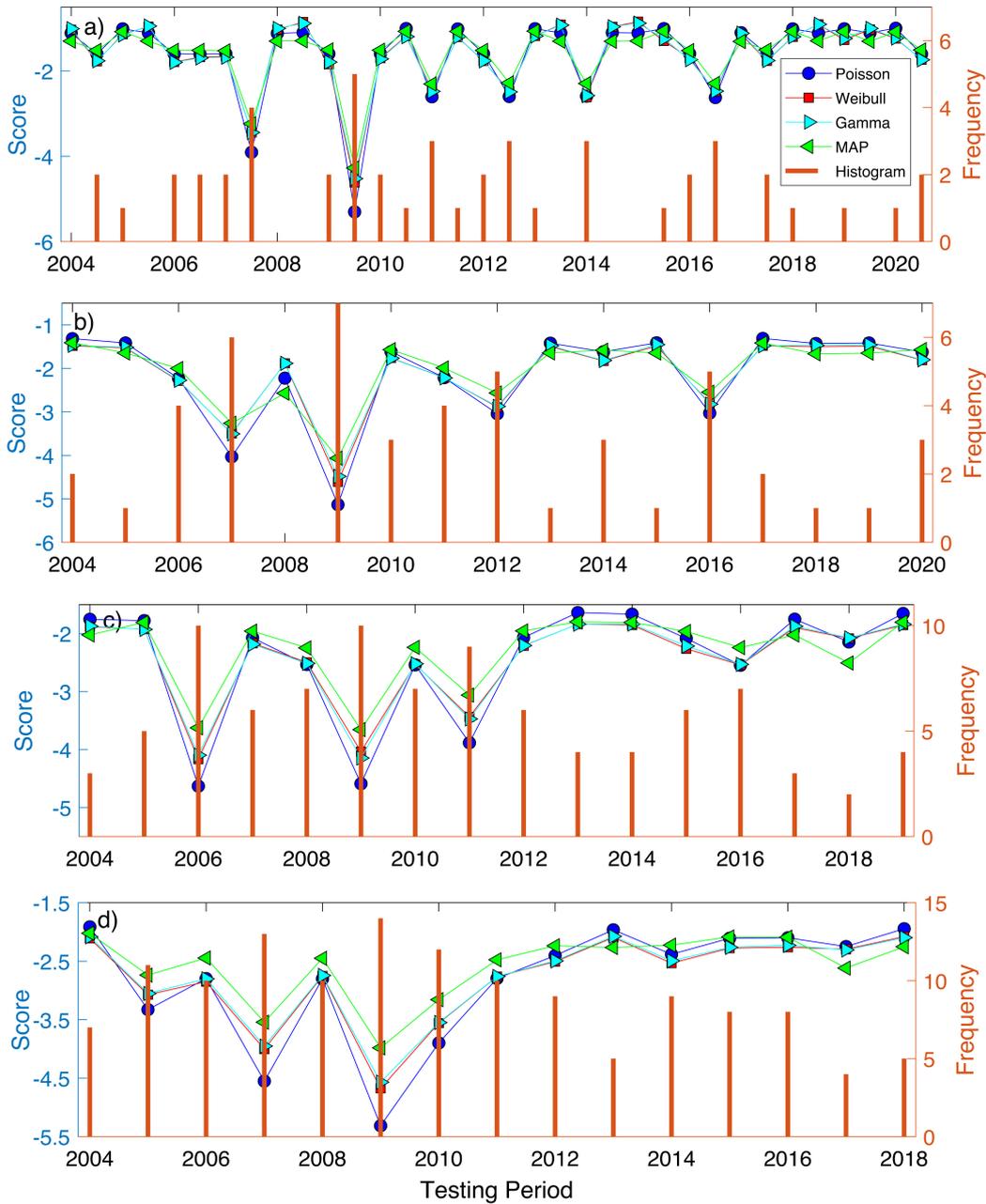


Figure 4.13: Scores of the forecasting models defined by Equation 4.1 for (a) 6-months, (b) 1-year, (c) 2-year and (d) 3-year long testing periods over the 17-year testing interval for events with  $M \geq 7.6$ . Different colors and shapes correspond to different testing models. Orange vertical bars show the observed number of events during each testing period at the starting time of each interval. Values closer to zero indicate better performance.

For the MAP model, all the simulations are initiated at the starting time

of each testing period based on the evaluated state probability, following the procedure described in Section 2.5. In this way, we use the full potential of the model, as it has been illustrated in Figure 4.9. The forecasting probabilities,  $p_{i,k}$ , are estimated for each interval  $T_i = [t_i^*, t_i^* + \delta)$ ,  $\delta = 0.5, 1, 2, 3$  years, and each forecasting model. If the MAP model provides more accurate forecasts for the number of earthquakes in the testing period, then the difference in the logarithmic scores,  $S_i$ , should be positive. We observe in Figure 4.13 that the MAP model achieves higher scores than the other models for most of the testing bins,  $T_i$ , independently of their duration. Especially in periods with high frequency the differences are larger, like during 2007-2008 and 2009-2010 in Figure 4.13b or 2006-2008 and 2009-2011 in Figure 4.13c. This is probably due to the large variance that characterizes the counting process of the MAP model. Conversely, in periods with low frequencies the renewal models are superior, especially in testing periods with zero observed events. The information that the renewal models consider for the elapsed time since the last event is more critical for testing periods with few events, however, they produce less significant results as the number of observed events increases. The Weibull and Gamma scores for consecutive periods with zero observed events are increasing like in  $T_9 = (2008, 2008.5]$  and  $T_{10} = (2008.5, 2009]$  in Figure 4.13a with  $S_9 = -1.02, -1.00$  and  $S_{10} = -0.85, -0.88$  for Weibull and Gamma models, respectively. Increasing the duration of the testing interval  $\delta$ , the differences between the MAP and the renewal models are getting higher.

We expand the comparison using the differences among the total forecasting scores of each model. This is essentially the difference in the sum of the scores over all the testing periods for each model given in Equation 4.2. As it can be seen in Table 4.3, the superiority of the MAP model is increasing with longer testing intervals for earthquakes with  $M \geq 7.6$ . It is also clear that MAP model performs better than the Poisson process for all testing intervals, with  $D = 1.34, 1.68, 2.61, 3.96$ , respectively. The differences with Weibull ( $D = 1.44, 1.25, 2.45, 3.24$ ) and Gamma ( $D = 1.27, 0.98, 2.43, 2.88$ ) models

CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING  
LARGE EARTHQUAKES NUMBER

are also in favor of the MAP model for all testing intervals (Table 4.3). The Poisson model exhibits the worst performance among the compared models.

Table 4.3: Logarithmic scores of the MAP, Weibull, Gamma and Poisson models for the data sets with  $M \geq 7.0$ ,  $7.7$  and  $M \geq 7.8$ , respectively. The differences,  $D$ , are given in parenthesis.

Model	$\delta$	$M_{thr} = 7.0$	$M_{thr} = 7.6$	$M_{thr} = 7.7$	$M_{thr} = 7.8$
$S_{MAP}$	6 month	-79.24	-53.33	-47.63	-44.55
	1 year	-49.19	-34.78	-31.11	-29.02
	2 years	-53.07	-36.72	-35.12	-34.79
	3 years	-55.21	-38.58	-37.98	-38.89
$S_{pois}$	6 month	-80.50 (1.26)	-54.6753 (1.34)	-48.80 (1.16)	-46.15 (1.60)
	1 year	-50.80 (1.61)	-36.4761 (1.68)	-32.55 (1.43)	-31.02 (1.99)
	2 years	-55.37 (2.29)	-39.3415 (2.61)	-36.88 (1.75)	-37.43 (2.63)
	3 years	-58.55 (3.33)	-42.5438 (3.96)	-40.51 (2.53)	-41.99 (3.09)
$S_{wbl}$	6 month	-80.40 (1.15)	-54.7816 (1.44)	-49.16 (1.52)	-45.68 (1.12)
	1 year	-48.95 (-0.24)	-36.0391 (1.25)	-32.58 (1.46)	-30.81 (1.79)
	2 years	-52.99 (-0.08)	-39.1832 (2.45)	-37.08 (1.95)	-36.55 (1.75)
	3 years	-55.12 (-0.09)	-41.8295 (3.24)	-39.64 (1.66)	-39.72 (0.82)
$S_{gam}$	6 month	-79.93 (0.68)	-54.6078 (1.27)	-49.14 (1.50)	-45.78 (1.23)
	1 year	-48.96 (-0.22)	-35.7732 (0.98)	-32.39 (1.28)	-30.76 (1.73)
	2 years	-53.08 (0.01)	-39.1589 (2.43)	-36.83 (1.70)	-36.52 (1.73)
	3 years	-55.01 (-0.20)	-41.4690 (2.88)	-39.31 (1.33)	-39.75 (0.85)

Similar results are derived for the data sets with  $M \geq 7.7$  and  $M \geq 7.8$  that are shown in Table 4.3. For the data set with the lowest magnitude cutoff,  $M_{thr} = 7.0$ , we observe that the differences in Gamma ( $D = 0.68, -0.22, 0.01, -0.20$ ) and Weibull ( $D = 1.15, -0.24, -0.08, -0.09$ ) models fluctuate around zero, without permitting us to make an inference on the superiority of one model from the other. The decrease of the magnitude threshold might lead to the dominance of the short-term clustering, so more complicated MAP models might be needed, i.e., more states or different assumptions for the existence of the idle state.

Next, we check whether the difference in the scores between the MAP and the Poisson, Gamma, and Weibull, models,  $D = S_{MAP} - S_{ref}$ , is statistically significant. Therefore, we generate 10,000 simulated catalogs, namely 10,000 testing periods, assuming that these events are generated according to a ref-

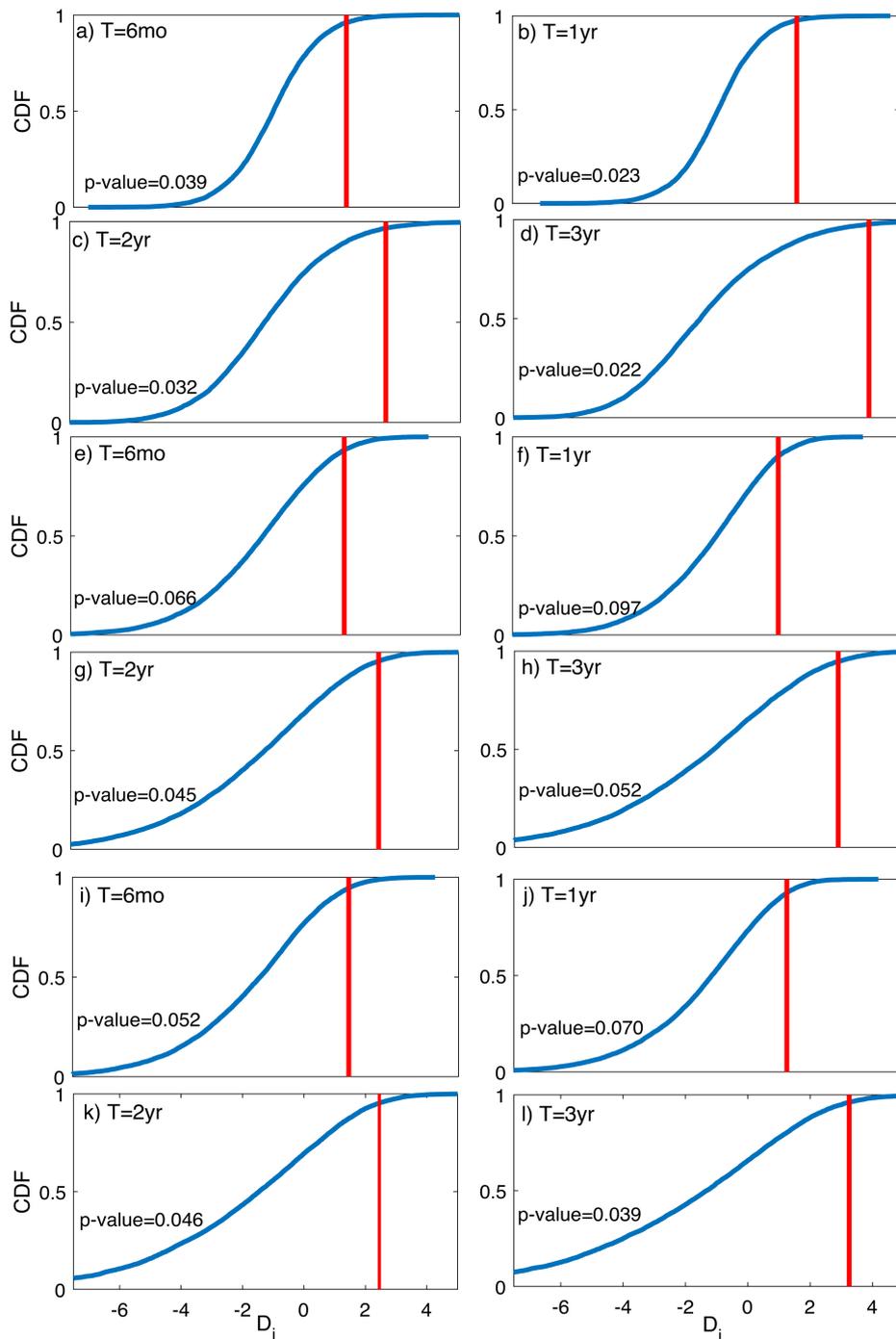


Figure 4.14: Cumulative distribution of the 10,000  $D_i$  values calculated from Equation 4.2 and from simulated samples assuming that the earthquakes are generated according to a Poisson a-d), Gamma e-h) and Weibull i-l) distribution, for events with  $M_{thr} = 7.6$  and 6-month, 1-year, 2-year and 3-year testing periods, respectively.

## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

---

reference model, in our case the ones mentioned above, for the data set with  $M_{thr} = 7.6$ . For each sample we calculate the two logarithmic scores as before and then their corresponding differences,  $D_i$ , with  $i = 1, \dots, 10,000$ . The observed  $D$  is indicative of a statistically better performance of the model with respect to the reference model if the associated  $p$ -value is low enough. Figure 4.14 shows the cumulative distribution of the differences  $D_i$  between the MAP model and the Poisson, Gamma and Weibull models according to which 10,000 samples were generated. The  $p$ -values indicate that the scoring differences are statistically significant, except from few cases where the  $p$ -values are larger than 0.05 but still less than 0.1 which provide evidence against the null hypothesis. When Gamma is used as a reference model the  $p$ -values are  $p = 0.066$  and  $p = 0.097$  for  $\delta = 6$  months and  $\delta = 1$  year (Figures 4.14e,f), respectively, indicating that the probability to observe the differences  $D = 1.27$  and  $D = 0.98$  (red vertical lines in Figures 4.14e,f and Table 3) when data is generated from a Gamma renewal model are still relatively low. For longer testing intervals the probability to observe the differences  $D = 2.43$  and  $D = 2.88$  for  $\delta = 2$  and  $\delta = 3$  years, respectively, are even lower ( $p = 0.045$  and  $p = 0.052$  in Figures 4.14g,h, respectively). For Poisson as a reference model the observed differences ( $D = 1.34, 1.68, 2.61, 3.96$  given in Table 4.13) are statistically significant for all testing intervals with corresponding  $p$ -values equal to  $p = 0.039, 0.023, 0.032, 0.022$  (Figures 4.14a,b,c,d). Finally, when we consider Weibull as a reference model the observed differences between MAP and Weibull ( $D = 1.44, 1.25, 2.45, 3.24$  given in Table 4.13) seem to be significant as indicated by the low  $p$ -values almost in all cases ( $p = 0.052, 0.070, 0.046, 0.039$  for  $\delta = 6$  months and  $\delta = 1, 2, 3$  years, respectively in Figures 4.14i,j,k,l).

Same observations are made for the two data sets with  $M \geq 7.7, 7.8$ . In particular, Figure 4.15 shows the the cumulative distribution of the differences,  $D_i$ , between the MAP and Gamma models, which exhibit the smallest differences among the tested models (see Table 4.13). Apart from Figure

CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

4.15h for  $\delta = 3$  years testing interval, in all other cases, the probability to observe the initial differences ( $D = 1.50, 1.28, 1.70, 1.33$  for  $M_{thr} = 7.7$  and  $D = 1.23, 1.73, 1.73, 0.85$  for  $M_{thr} = 7.8$ ) is very low, in particular,  $p < 0.1$ . The same or even lower  $p$ -values are derived for the other models.

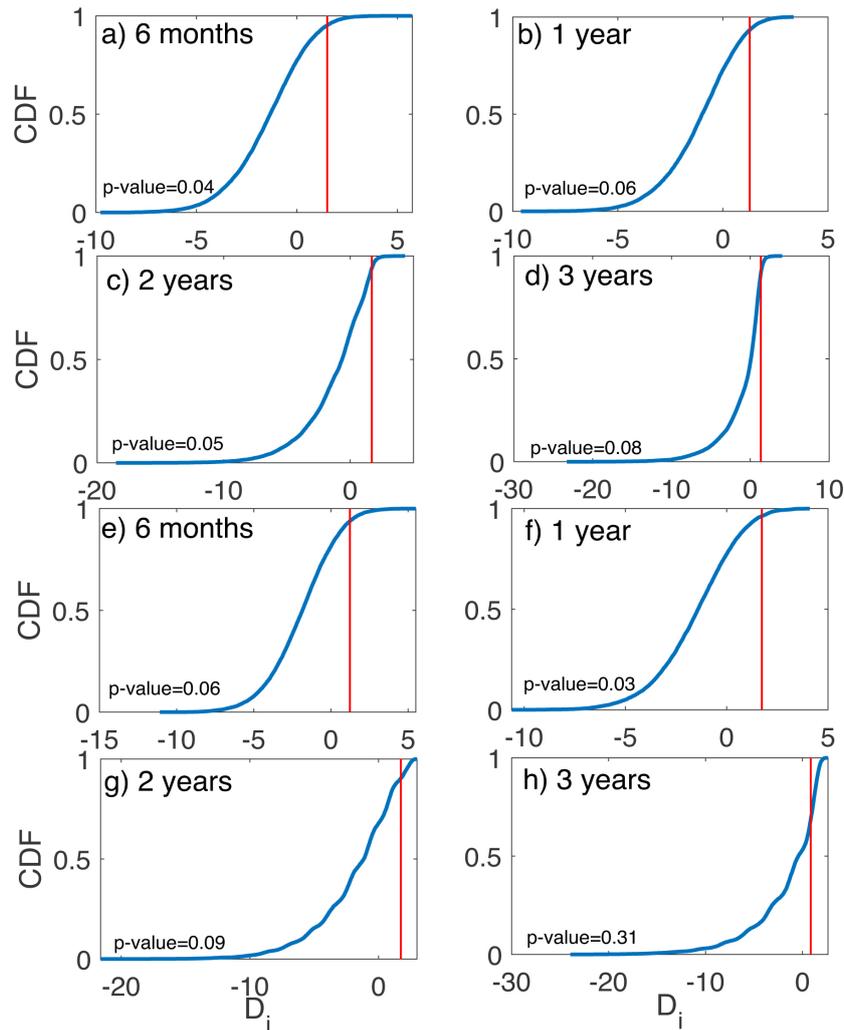


Figure 4.15: Cumulative distribution of the 10,000  $D_i$  values calculated from Equation (4.2) and from simulated samples assuming that the earthquakes are generated according to a Gamma distribution, for events with  $M_{thr} = 7.7$  a-d) and  $M_{thr} = 7.8$  e-h), respectively.

#### 4.4.3.1. Comparison with the ETAS model

Gamma and Weibull renewal processes proved insufficient especially for capturing the short-term clustering of seismicity. For this reason, the temporal ETAS model is also fitted to the learning period, which is a model known for its efficiency in representing aftershock sequences. Figure 4.16 shows the logarithmic scores of ETAS and MAP models for earthquakes with  $M \geq 7.6$  following the procedure described in Section 4.3.2. We find that the scores of the MAP model are higher than the ETAS ones in most of the intervals, with some exceptions again in testing intervals with few events. Similar results are obtained for thresholds  $M_{thr} = 7.7, 7.8$ , where the total scores over all the testing periods shown in Table 4.4 indicate the superiority of the MAP model. An additional background rate applied to the intensity function of the ETAS model might be necessary for increasing its forecasting skill. Only for the dataset with  $M_{thr} = 7.0$ , the ETAS model seems to perform better, where MAP yields similar results when compared to the Gamma and Weibull models (Table 4.13). The effect of aftershocks and foreshocks is enhanced when we consider this magnitude cutoff, so it might be more appropriate to consider the initial MAP model or increase the states of the current modified MAP model that is used in the study.

Table 4.4: Logarithmic scores of the MAP and ETAS models. The differences,  $D$ , are given in parenthesis.

Model	$\delta$	$M_{thr} = 7.0$	$M_{thr} = 7.6$	$M_{thr} = 7.7$	$M_{thr} = 7.8$
$S_{MAP}$	6 month	-79.24	-53.33	-47.63	-44.55
	1 year	-49.19	-34.78	-31.11	-29.02
	2 years	-53.07	-36.72	-35.12	-34.79
	3 years	-55.21	-38.58	-37.98	-38.89
$S_{etas}$	6 month	-79.31 (0.07)	-54.72 (1.38)	-48.69 (1.05)	-46.32 (1.77)
	1 year	-48.61 (-0.57)	-36.19 (1.40)	-32.14 (1.03)	-31.14 (2.11)
	2 years	-51.45 (-1.62)	-39.42 (2.69)	-36.66 (1.54)	-37.69 (2.90)
	3 years	-52.44 (-2.76)	-42.62 (4.04)	-39.51 (1.53)	-42.27 (3.38)

CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING  
LARGE EARTHQUAKES NUMBER

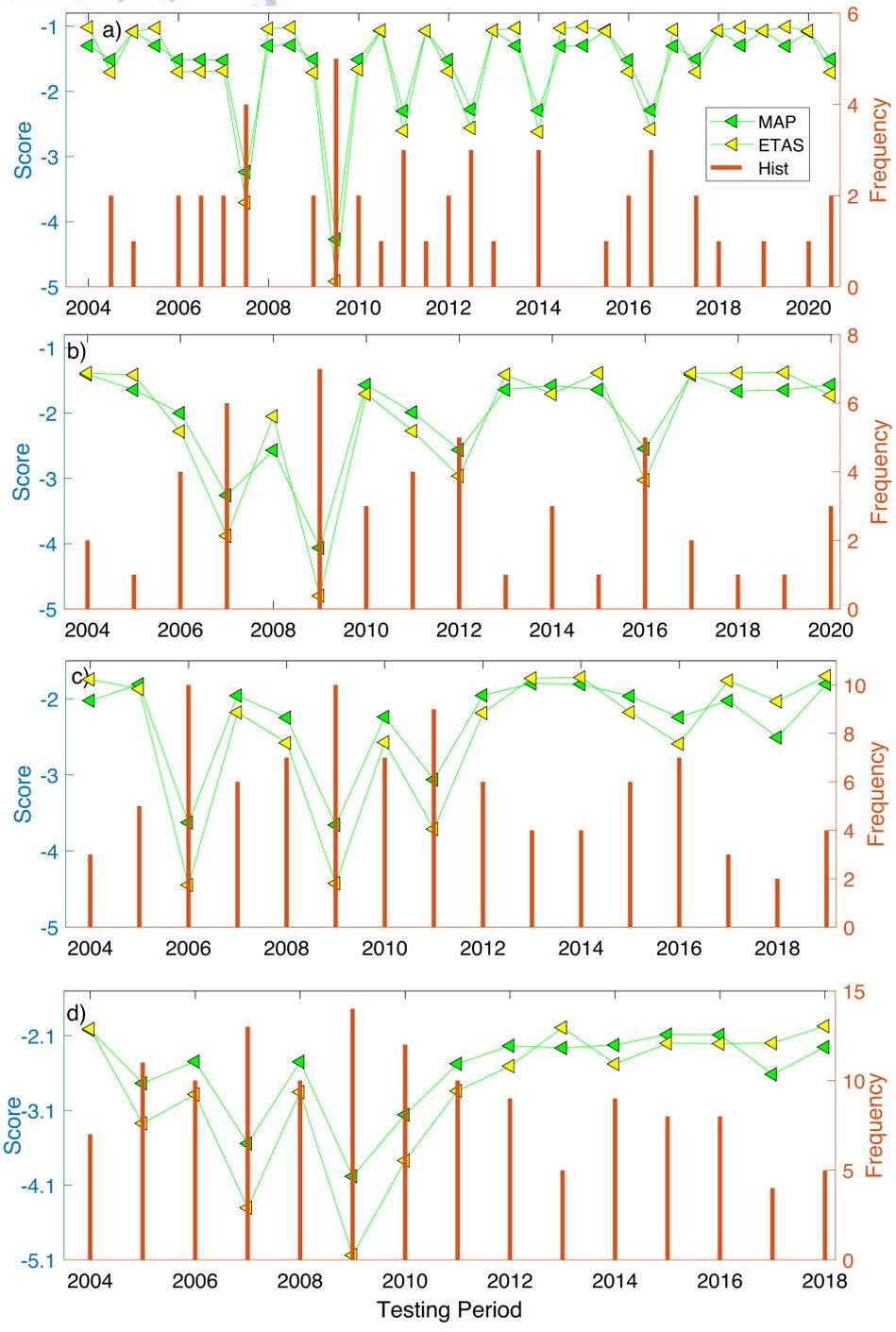


Figure 4.16: Scores of the forecasting models for (a) 6-months, (b) 1-year, (c) 2-year and (d) 3-year long testing periods over the 17-year testing interval for earthquakes with  $M \geq 7.6$ . Different colors and shapes correspond to different testing models. Orange vertical bars show the observed number of events during each of the testing periods at the starting time of each interval. Values closer to zero indicate better performance.

#### 4.4.3.2. Comparison with the initial MAP

We showed in Section 4.4.1 that when we consider the initial MAP, i.e., without implementing the two-step procedure, the estimated seismicity rates cannot discriminate the long periods of seismic quiescence from periods of moderate seismicity. State 2 and state 3 (yellow and orange color in Figure 4.5) correspond to high seismicity rates for all data sets and are associated with the occurrence of events in temporal proximity, whereas the main seismic activity is approximated by a stable rate associated with state 1 (red color in Figure 4.5). Nevertheless, we computed the forecasting efficiency of the initial MAP and the one proposed in this study with the idle state. Figure 4.17 shows that the initial MAP yields a worse performance in terms of the forecasting scores over almost all testing intervals for earthquakes with  $M \geq 7.6$ . Same behavior is observed for the other large magnitude thresholds (Table 4.5). Our proposed two-step estimation procedure for the MAP model contributes substantially to the forecasting of the earthquake occurrences, especially for the higher magnitude thresholds ( $M_{thr} = 7.6, 7.7, 7.8$ ). Worth mentioning is the slight superiority of the initial MAP compared to the MAP with the idle state for earthquakes with  $M \geq 7.0$ . As we mentioned earlier, short-term clustering effect seems to dominate as we decrease the magnitude cutoff, so it might not be appropriate to consider the two-step procedure for lower thresholds.

Table 4.5: Logarithmic scores of the modified MAP,  $MAP_{idle}$ , and the initial one,  $MAP_{init}$ . The differences,  $D$ , are given in parenthesis.

Model	$\delta$	$M_{thr} = 7.0$	$M_{thr} = 7.6$	$M_{thr} = 7.7$	$M_{thr} = 7.8$
$MAP_{idle}$	6 month	-79.24	-53.33	-47.63	-44.55
	1 year	-49.19	-34.78	-31.11	-29.02
	2 years	-53.07	-36.72	-35.12	-34.79
	3 years	-55.21	-38.58	-37.98	-38.89
$MAP_{init}$	6 month	-78.35 (-0.88)	-54.7970 (1.46)	-49.0052 (1.36)	-46.02 (1.47)
	1 year	-48.74 (-0.44)	-35.94 (1.15)	-32.20 (1.08)	-30.78 (1.75)
	2 years	-53.06 (-0.008)	-39.18 (2.45)	-36.81 (1.68)	-37.07 (2.28)
	3 years	-55.13 (-0.08)	-42.14 (3.56)	-39.97 (1.99)	-41.06 (2.16)

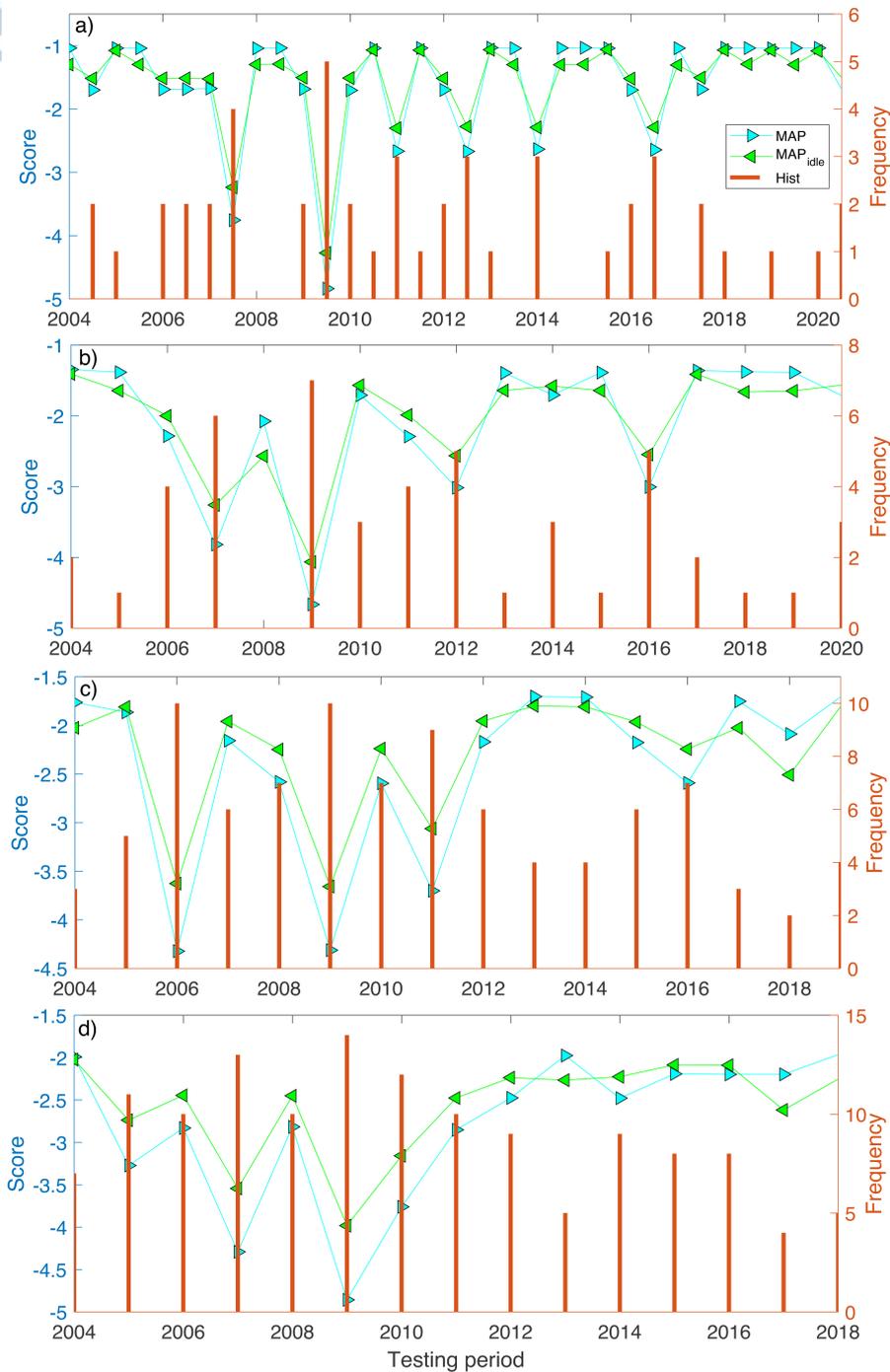


Figure 4.17: Scores of the forecasting models for (a) 6-months, (b) 1-year, (c) 2-year and (d) 3-year long testing periods over the 17-year testing interval for earthquakes with  $M \geq 7.6$ . Different colors and shapes correspond to different testing models. Orange vertical bars show the observed number of events during each of the testing periods at the starting time of each interval. Values closer to zero indicate better performance.

## 4.5 π. Discussion

In this study, we provided an advancement towards the long-term modeling of large earthquake occurrences by contributing to the forecasting of their occurrence number. Towards this direction, we followed the [Nandan et al. \(2019\)](#) suggestion that "...the idea of embracing the full distribution of earthquake numbers should be extended to earthquake forecasting models of other types, such that their true forecasting potential is revealed." We introduced a two-step modeling procedure for the MAP model, that reveals the non-stationary characteristics of the temporal distribution of the large earthquakes ( $M \geq 7.6$ ,  $M \geq 7.7$  and  $M \geq 7.8$ ) in circum-Pacific belt taking advantage of our empirical observations on the data sets and we conducted a catalog-based pseudo-prospective experiment considering the full earthquakes distribution.

There are some efforts to create a link between a plausible physical mechanism and the worldwide temporal clustering of large events under the concept of seismic cycles synchronization. Dynamic stress interactions between distant faults can advance or delay their respective seismic cycles, leading to synchronization, or in other words to the production of temporal clusters of events ([Sammis and Smith, 2013](#)). [Bendick and Bilham \(2017\)](#) statistically quantified the worldwide synchronization of earthquakes with  $M \geq 7.0$  since 1900 through topological networks and time series analyses and [Bendick and Mencin \(2020\)](#) underpinned the occurrence of large events as a time-dependent process that depends on previous seismicity with similar recurrence intervals and occurrence times and therefore a degree of predictability might exist. Our speculation is that earthquake occurrences close in time among different regions can be superimposed producing elevated levels of seismic activity for the large earthquakes ( $M \geq 7.6, 7.7, 7.8$ ) in a global scale for irregular periods varying among them from a couple of years to a couple of decades. In this case, state 1 of the MAP model might capture long-term periods of increased seismicity that could comprise independent synchronizations, different groups

of faults with aligned seismic cycles, associated to the occurrence of some of the largest instrumentally recorded events, since 12 out of the 15 events with  $M \geq 8.5$  occurred in periods of increased seismicity for all magnitude thresholds. Hence, non-stationarity might characterize their temporal distribution, where persistent periods of increased seismic activity alternate with long irregular intervals of seismic quiescence. Finally, the regional spatiotemporal clustering is expressed through state 2 of the model for great events along the Sumatra, Japan, Kuril, Tonga, Solomon Islands and New Hebrides subduction zones (Figure 4.12).

The investigation of the temporal patterns of large earthquake occurrences in circum-Pacific belt revealed the existence of non-regular periods of increased seismicity, including short-term seismicity clusters, and long-term changes of the seismicity rate. The short-term earthquake clustering in close distances can be to a degree attributed to fault interactions due to static stress (Parsons and Velasco, 2011; Scholz, 2010) and even though it can be related to the existence of large aftershocks or foreshocks, they should not be removed by a declustering algorithm since they are as important as main shocks in terms of seismic hazard. We believe that the incorporation of non-stationarity for the forecasting of the large earthquakes number is significant for earthquake hazard assessment because estimates for future occurrence probabilities are based on whether it is considered a period of increased seismicity or a relative quiescence according to the earthquake history.

## 4.6 Conclusions

In this Chapter, we established a two-step estimation procedure via a Markovian Arrival Process to approximate the temporal distribution of large earthquakes in circum-Pacific belt. Long inter-event times are considered extreme events due to their rarity and they are modeled separately from the rest of the seismicity. In this way, non-stationary characteristics are revealed for



## CHAPTER 4. MARKOVIAN ARRIVAL PROCESS FOR FORECASTING LARGE EARTHQUAKES NUMBER

---

large earthquake occurrences that are persistent for magnitude thresholds  $M_{thr} = 7.6$ ,  $M_{thr} = 7.7$  and  $M_{thr} = 7.8$ .

Our findings imply that: (1) there is evidence for non-stationarity in the observed seismicity with  $M_{thr} = 7.6$ ,  $7.7$  and  $M_{thr} = 7.8$ , although less pronounced for  $M \geq 7.0$ , that due to the presence of short-time events the MAP model cannot approximate well. After removing the short-term clustered events from the initial data sets with  $M \geq 7.0$  and  $M \geq 7.6$ , long-term changes in the background seismicity rate are revealed by the MAP model (see Section 4.4.1), and two common, for both data sets, long seismic quiescence periods exist during 1948-1966 and 1981-1991 (red color in Figure 4.5). (2) Incorporating an idle state to the parameter set of the MAP model that approximates separately the long inter-event times we revealed two distinct patterns that are preserved with small variations for the data sets with the higher magnitude thresholds ( $M_{thr} = 7.6$ ,  $M_{thr} = 7.7$  and  $M_{thr} = 7.8$ ), i.e. long-term interactions varying from a couple of years to a couple of decades and intense spatiotemporal clustering (Figures, 4.9, 4.10 and 4.12). (3) The pseudo-prospective forecasting experiments indicate that the modified MAP model outperforms the Poisson, Weibull and Gamma models during all testing intervals and for the three higher magnitude cutoffs ( $M_{thr} = 7.6$ ,  $7.7$  and  $M_{thr} = 7.8$ ) of the data sets (Tables 4.3 and Figures 4.13, 4.14, 4.15). It yields also better performance over the temporal ETAS model during all testing intervals. However, the two-step procedure might not be appropriate for lower magnitude thresholds since the MAP model returns comparable or slightly worse results against the Gamma and ETAS models for  $M_{thr} = 7.0$  (Figure 4.16 and Tables 4.3, 4.4) and the initial MAP yields better scores compared to the MAP with the idle state for earthquakes with  $M \geq 7.0$  (Table 4.5).





## Chapter 5

# Concluding Remarks

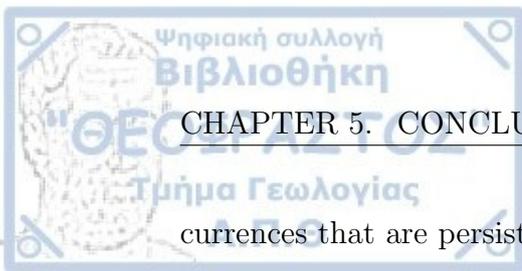
Earthquake clustering is an essential feature of seismicity that contributes to the understanding of the long-term trends of seismicity and to reveal its short-scale dynamics. In the present thesis, we have studied the short and long-term properties of the temporal distribution of seismicity through the use of stochastic modeling and explored the difficulties in addressing both types of clustering.

Our first contribution, is the proposal of the Markovian Arrival Process for the approximation of the earthquakes temporal distribution. Towards this direction, we constructed a grid-based procedure for the selection of the initial values of the *EM* algorithm and developed a parallel framework for its implementation. In this way, we improved the chances to converge to the maximum value of the log-likelihood function and reduced the required computation time. We also proposed the use of a local decoding algorithm for the evaluation of the most probable path of hidden states for the model. Its stability is verified on simulated data sets where the sequence of the hidden states is known. This is extremely important, because it provides us with the seismicity rate of the counting process of the model at each time of the study period. In this way, the MAP model can be used a change point tool for the detection of changes in seismicity rate.

In Chapter 3 we verified the efficiency of the MAP model to detect seismicity rate changes on well studied seismic sequences of the Corinth Gulf earthquake catalog during 1964-2017 for events with  $M \geq 4.5$ . We showed that the identified seismicity rate changes are mainly related either to main shock-aftershock sequences or earthquake swarms. Based on this observation, we established a new clustering procedure, which we called MAP-DBSCAN. The method proved efficient on detecting the clusters of a simulated ETAS catalog where the links among the events are known. We used the Jaccard index as a validation metric, which we believe is an appropriate tool in performance studies for earthquake clustering. Moreover, we showed the competitiveness of the MAP-DBSCAN procedure against well-known clustering algorithms, as in most cases, exhibits better results.

We applied the method to three major seismic zones of Greece and investigated their clustering properties. The detected seismic clusters in the Corinth Gulf, Central Ionian Islands and North Aegean Sea during 2012–2019 for events with  $M \geq 1.5$ ,  $M \geq 2.2$  and  $M \geq 2.1$ , respectively, are concentrated to the main seismotectonic structures of the areas. We examined the regional variability among the three areas by inverting the generic ETAS parameters with a stacking procedure. The aftershock productivity is extremely high in Central Ionian Islands, where main shock-aftershocks are dominant, whereas Corinth Gulf is characterized by low productivity values and high background rates due to the existence of swarm activity. We also inferred sequence-specific parameters of the temporal ETAS model for clusters with  $N \geq 30$ . Low productivity values and high background rates are related to sequences that occurred in areas with high pore-pressure due to fluids migration.

In Chapter 4, we proposed the use of the Markovian Arrival Process for the modeling of the long-term properties of large earthquakes in circum-Pacific belt. To confront the difficulties in modeling both the short-time values and long quiescence periods we established a two-step estimation procedure. In this way, non-stationary characteristics are revealed for large earthquake oc-



## CHAPTER 5. CONCLUDING REMARKS

---

currences that are persistent for high magnitude thresholds ( $M_{thr} \geq 7.6$ ).

We showed that the existence of short-term seismicity complicates the study of the long-term variations in the temporal distribution of large earthquakes. After removing the short-term clustered events from the initial data sets, we provided evidence for long-term changes in the background seismicity rate. Incorporating an idle state to the parameter set of the MAP model that approximates separately the long inter-event times we revealed two distinct patterns that are preserved with small variations for the all the data sets with high magnitude thresholds ( $M_{thr} \geq 7.6$ ), i.e. long-term interactions varying from a couple of years to a couple of decades and intense spatiotemporal clustering. The pseudo-prospective forecasting experiments that we performed with the modified MAP model show better results than the Poisson, Weibull and Gamma models and the temporal ETAS model for multiple testing intervals and  $M_{thr} \geq 7.6$ . For lower magnitude thresholds the modified MAP model returns comparable or slightly worse results against the Gamma and ETAS models for  $M_{thr} = 7.0$ .



# A

## Appendix A

### A.1 ETAS formulation

The ETAS model is a stochastic point process that incorporates the basic empirical laws of seismicity in space and time. [Ogata \(1988\)](#) has established the temporal ETAS model, that is essentially a self-exciting Hawkes process (see [Hawkes \(1971\)](#)) that is defined by the conditional intensity function,

$$\lambda(t) = \lambda(t/H_t) = \lim_{dt \rightarrow 0} \frac{P_{dt}(t/H_t)}{dt} = \mu + \sum_{j:t_j < t} h(t - t_j), \quad (\text{A.1})$$

where  $P_{dt}$  gives the probability of an earthquake to occur at an infinitesimal time interval  $(t, t + dt)$ , conditional on the history of the process,  $H_t = \{(t_j, m_j) : t_j < t\}$ , with  $t_j$  the occurrence times up to time  $t$  and corresponding magnitudes  $m_j$ .

The first part of Equation A.1,  $\mu$ , is the background rate (migration rate) that is stationary in time and is based on the constant tectonic loading that produces independent earthquakes (mother events). The second part is the triggering function, that expresses the contribution of past events with  $t_j < t$  to the occurrence rate of a new event at time  $t$ . This is based on the main assumption of ETAS model that each event is possible to trigger a future

event or in terms of epidemic models that both mother events and daughters can generate their own daughters.

Ogata (1988) adopted the two following main empirical laws that express the time-dependent seismicity in the triggering function,  $h(t-t_j) = k(m_j)g(t-t_j)$ :

- the productivity law (Utsu (1971) pp 420-427),

$$k(m_j) = Ke^{a(m_j-m_c)}, \quad (\text{A.2})$$

which gives the number of aftershocks with  $M \geq m_c$  triggered by a mother event with magnitude  $m_j$ ;

- the modified Omori law,

$$g(t-t_j) = (p-1)c^{(p-1)}(t-t_j+c)^{-p}, \quad (\text{A.3})$$

with  $p > 1$ , that describes the temporal decay of aftershocks.

The  $a$  parameter of the productivity law, given by Equation A.2, scales the effect of an event with magnitude  $m_j$  to the triggering of first generation aftershocks. A large value of  $a$  means that the proportion of triggered events from large earthquakes is higher than from small ones. This is easily understood from the equation  $k(M_1)/k(M_2) = e^{a(M_1-M_2)}$  between two events with magnitudes  $M_1$  and  $M_2$ , respectively. For instance, when  $a = 2.0$  then the fraction between the produced aftershocks of the events with  $M_1 = 7.5$  and  $M_2 = 5.5$ , respectively, is almost 55 (the first event will produce 55 times more aftershocks than the second), whereas when  $a = 1.0$  the ratio of the same events would be almost 8. It has been observed that small  $a$  values characterize swarm type sequences where secondary bursts of seismic activity occur, whereas higher  $a$  values are usually observed in earthquake catalogs that are dominated by main shock-aftershock sequences (Hainzl and Ogata, 2005). Parameter  $K$  gives the expected aftershock productivity independently of the

main shock magnitude and is depending on the magnitude cutoff when  $a \neq \beta$  (Seif et al., 2017), where  $\beta = b \cdot \ln 10$ .

The modified Omori law is a power-law function that gives the decay of aftershocks with time. Some studies suggest that parameter  $c$  reflects the incompleteness in the early stage of the aftershock sequences (Kagan and Knopoff, 1981) and is also depending on the cutoff magnitude (Seif et al., 2017). More interpretations can be derived for exponent  $p$  that describes the decay rate of the aftershock sequence. Higher  $p$  values indicate fast diminishing sequences whereas as  $p$  decreases the duration of an aftershock sequence is elongating. Due to normalization of function,  $g(t)$ ,  $p$  should always be  $p > 1$ .

Following the formulation of the temporal ETAS model, Ogata (1998) extended the functional form of the conditional intensity function,  $\lambda(t)$ , embedding the space component of seismicity. It takes the following form,

$$\lambda(t, \mathbf{x} = (x, y)/H_t) = \lim_{dt, dx, dy \rightarrow 0} \frac{P_{dt, dx, dy}(t/H_t)}{dt dx dy} = \mu(\mathbf{x}) + \sum_{j: t_j < t} h(t - t_j, \mathbf{x} - \mathbf{x}_j), \quad (\text{A.4})$$

where  $P_{dt, dx, dy}$  gives the probability of an earthquake occurrence at an infinitesimal time interval  $(t, t + dt)$  and in a small region  $[x, x + dx) \times [y, y + dy)$ . The background rate,  $\mu(\mathbf{x})$ , is stationary in time and heterogeneous in space due to the concentration of events around the faults and the triggering function consists of three terms,  $h(t - t_j, x - x_j, y - y_j) = k(m_j) \cdot g(t - t_j) \cdot f(\mathbf{x} - \mathbf{x}_j)$ . The last component of the equation is the spatial distribution of aftershocks,

$$f(\mathbf{x} - \mathbf{x}_j/m_j) = \frac{q - 1}{\pi d(m_j)^{q-1}} [||\mathbf{x} - \mathbf{x}_j||_2^2 + d(m_j)]^{-q}, \quad (\text{A.5})$$

with  $q > 1$ , and  $d(m_j) = d_0 10^{\gamma(m_j - m_c)}$ , considering an isotropic distribution of aftershocks around the main shock. Parameter  $q$  is the power law exponent of the spatial distribution and gives the decay of aftershocks in space, whereas  $d_0$  expresses the spatial spreading of aftershocks. Small values of  $d_0$  indicate high concentration and vice versa. Parameter  $\gamma$  scales the aftershock spreading

with the main shock magnitude  $m_j$ .

## A.2 Estimation procedure

The parameter set of the space-time ETAS model consists of 8 variables,  $\boldsymbol{\theta} = \{a, K, p, c, d, \gamma, q, \mu\}$ . The most common estimation method is the maximization of the log-likelihood function,  $L = L(\boldsymbol{\theta}/Tr)$ , with  $Tr = \{(t_i, x_i, y_i, m_i), i = 1, \dots, N\}$  the events that occurred in the spatio-temporal region  $A = \Sigma \times [t_{in}, t_f]$  with  $M \geq m_c$  and is given by

$$LL = \sum_{i \in A} \log \lambda(t, \mathbf{x}) - \iint_{\Sigma} \int_{t_{in}}^T \lambda(t, \mathbf{x}) dt d\Sigma. \quad (A.6)$$

Concerning the evaluation of  $\lambda(t, \mathbf{x})$ , it is shown that the missing links between events inside the target region  $A$  and below the magnitude threshold, before the start and outside the spatial boundaries of the data set influence the final  $LL$  value and as a consequence the inverted estimated parameters (Wang et al., 2010). So, events in the broader region  $\Sigma_0$  and time interval  $[t_0, T]$ , with  $t_0 < t_{in}$  should be included in the evaluation of,  $\lambda(t, \mathbf{x})$ .

For the maximization of the  $LL$  function we implement an iterative procedure following Lippiello et al. (2014). An initial parameter set,  $\boldsymbol{\theta}_0$ , is chosen based on the seismic features of the corresponding study area. Then, at each iteration step, ( $r$ ), we update the model parameters by adding a random factor so  $\theta_k^{(r)} = \theta_k^{(r-1)} + u$ , for  $k = 1, \dots, 8$ . We compute the new log-likelihood value,  $LL^{(r)}$  and store the new parameters under the condition  $LL^{(r)} > LL^{(r-1)}$ . After some iterations, the logarithm converges and the algorithm stops. Essentially, this is a grid-based procedure, since we use a large number of iterations.

The simulation procedure that we use here is based on the branching structure of the ETAS model (Zhuang et al., 2004). In particular, each mother event generates its daughters (first generation), the daughters generate their own descendants (second generation), and so on.

For the simulation of an ETAS earthquake catalog, first, we need to generate the mother events based on the non-stationary in space background rate  $\mu(x, y)$ . The heterogeneity in space can be preserved from the spatial coordinates of the seismicity in the study area. We identify the background events in the original earthquake catalog using a declustering algorithm and then we generate  $N_{main}$  mother events from a Poisson distribution with mean value equal to the number of the identified background events. Their coordinates are sampled with replacement from the declustered catalog by adding a random factor. The occurrence times are simulated from a uniform distribution  $U(t_0, t_f)$ , where  $t_0$  and  $t_f$  are the starting and ending time of the simulated earthquake catalog, respectively.

The magnitudes are independent from the earthquakes' spatial and temporal distribution and follow the Gutenberg–Richter law truncated from the left at the completeness magnitude,  $m_c$ , and from the right at a maximum upper threshold,  $m_{max}$ . The functional form of their distribution is the following,

$$s(m) = \frac{\beta e^{-\beta m}}{(e^{-\beta m_c} - e^{-\beta m_{max}})}. \quad (\text{A.7})$$

Next, for each mother event  $i$  with  $i = 1, \dots, N_{main}$ , we simulate their aftershocks number from a Poisson distribution with expected rate equal to the productivity of the model,  $k(m_i) = K e^{a(m_i - m_c)}$ . Their occurrence times are sampled from the modified Omori law,  $g(t)$ , given by Equation (A.3), and the locations from the isotropic spatial distribution function,  $f(\mathbf{x})$ , given by Equation (A.5). Events outside the region  $\Sigma_0$ , time period  $(t_0, t_f)$  and below

the magnitude threshold,  $m_c$ , are neglected. For next-generation daughters the triggering step is repeated considering each daughter a mother event until there are no more generated events.

To avoid the boundary issue, i.e. triggering effect of events outside the target region and from an earlier period, we implement a simulation in a broader spatial and temporal area  $\Sigma_0 \times [t_0, t_f]$  and then consider events in the smaller target area  $\Sigma \times [t_{in}, t_f]$  with  $\Sigma \subseteq \Sigma_0$  and  $t_{in} > t_0$  for our study.



## B

# Appendix B

## B.1 Gardner and Knopoff algorithm

The procedure introduced by [Gardner and Knopoff \(1974\)](#) for the detection of aftershocks is based on specific magnitude dependent space-time windows. It is known as the window-based method, and it is one of the simplest forms of aftershock identification. For each earthquake with magnitude  $M$ , the subsequent events are assigned as aftershocks if they occur within a temporal window  $t(M)$  and a spatial interval  $d(M)$ , respectively. Foreshocks are treated as aftershocks when a larger earthquake occurs later in the sequence. The event is considered as an aftershock and the algorithm is repeated based on the largest magnitude.

We give in Equation (B.1) the functional form of the spatial and temporal windows suggested in [Gardner and Knopoff \(1974\)](#), which are denoted as GK1. Additionally, in Equations (B.2) and (B.3) we present alternative window parameter settings that can be found in [van Stiphout et al. \(2012\)](#). We denote them as GK2 and GK3, respectively.

$$d = 10^{0.1238*M+0.983} \text{ (km)} \text{ and } t = \left\{ \begin{array}{ll} 10^{0.032*M+2.7389} & M \geq 6.5 \\ 10^{0.5409*M-0.547} & M < 6.5 \end{array} \right\} \text{ days} \quad (\text{B.1})$$

$$d = e^{1.77 + \sqrt{0.037 + 1.02 * M}} \text{ (km) and } t = \left\{ \begin{array}{ll} 10^{2.8 + 0.024 * M} & M \geq 6.5 \\ e^{-3.95 + \sqrt{0.62 + 17.32 * M}} & M < 6.5 \end{array} \right. \text{ days} \quad (\text{B.2})$$

$$d = e^{-1.024 + 0.804 * M} \text{ (km) and } t = e^{-2.87 + 1.235 * M} \text{ days.} \quad (\text{B.3})$$

## B.2 Reasenberg Linked-Based algorithm

In the Reasenberg method (Reasenberg, 1985), an interaction zone among earthquakes is assumed that is modeled based on estimates of the stress redistribution for the spatial extent and on a probabilistic model, the Omori law, for the temporal extent, respectively. Any earthquake that occurs within the interaction zone of a prior earthquake is considered an aftershock and is included in the cluster. The parameters  $\tau_{min}$  and  $\tau_{max}$  correspond to the minimum and maximum elapsed time since the last event, in order to observe the next correlated earthquake at a certain probability,  $p_1$ . Additionally,  $x_{meff}$  denotes the minimum magnitude threshold for the earthquake catalog, whose value in the clusters is raised by a factor  $x_k$  of the largest earthquake,  $M$ , within, and is given by  $x_{meff} = x_{meff} + x_k * M$ . Finally, the parameter  $r_{fact}$  corresponds to the radii we adopt to consider linking a new event with the cluster. Overlapping clusters are merged.

## B.3 Nearest-Neighbor algorithm

The Nearest-Neighbor approach is based on the space-time-magnitude distance metric among two earthquakes given by Baiesi and Paczuski (2004):

$$\eta_{ij} = (t_j - t_i) r_{ij}^{d_f} 10^{-bm_i}, \quad (\text{B.4})$$

B. APPENDIX B

where  $r_{ij}$  is the epicentral distance between events  $i$  and  $j$ ,  $d_f$  is the spatial fractal dimension and  $b$  is the component of the Gutenberg–Richter distribution. Each event  $j$  is connected to its nearest neighbor  $i = \operatorname{argmin}_{i:t_j > t_i} \eta_{ij}$  if their distance,  $\eta_j$ , is lower than a predefined threshold  $\eta_0$ . The earthquake catalog is then partitioned on distinct clusters, each containing at least one event. The space and time distances between two events are normalized considering the logarithm of metric  $\eta_{ij}$  and they are given by

$$T_{ij} = (t_j - t_i)10^{-0.5bm_i}, \quad R_{ij} = r_{ij}^{d_f} 10^{-0.5bm_i},$$

so

$$\log \eta_{ij} = \log T_{ij} + \log R_{ij}. \quad (\text{B.5})$$

Zaliapin et al. (2008) showed that seismicity follows a bimodal distribution in relation to  $(T_{ij}, R_{ij})$ . One mode corresponds to the background seismicity and the other one is located in short space and time distances and is related to the correlated seismicity.

For the selection of the threshold value,  $\eta_0$ , the distribution of the logarithm of the nearest neighbor distance  $\eta^* = \{\eta_j\}_{j=1, \dots, N}$  is investigated, where  $N$  denotes the events number. It follows an 1D Gaussian distribution with two components, which is essentially a mixture model of two Gaussian densities with parameters  $N(\mu_1, \sigma_1)$ ,  $N(\mu_2, \sigma_2)$  and  $a_1$ ,  $a_2$  weights, respectively. The threshold value is defined as the intersection of the two modes.

# C

## Appendix C

### C.1 Additional figures of cluster analysis

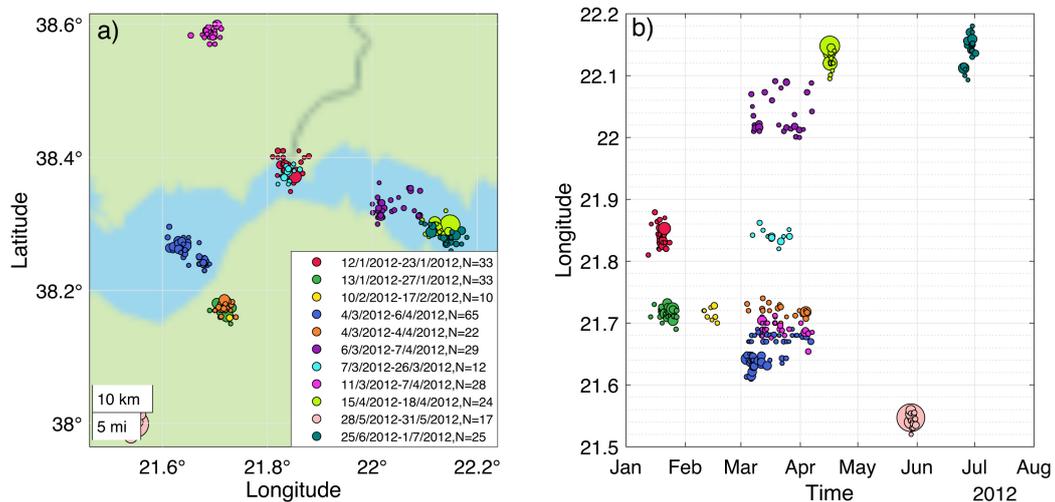


Figure C.1: a) Epicentral map of the main seismic clusters during the first semester of 2012. Three major clusters are occurred, the *C1*, *C2* and *C3* and eight smaller clusters with  $N \geq 10$  events. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

C. APPENDIX C

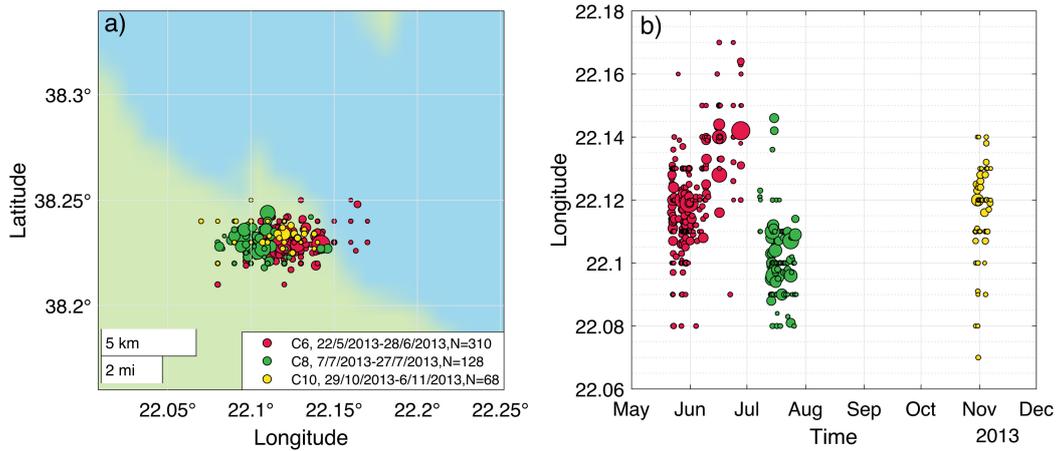


Figure C.2: a) Epicentral map of the 2013 Aigion swarm and subsequent sequences in the area with  $N \geq 10$  events. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

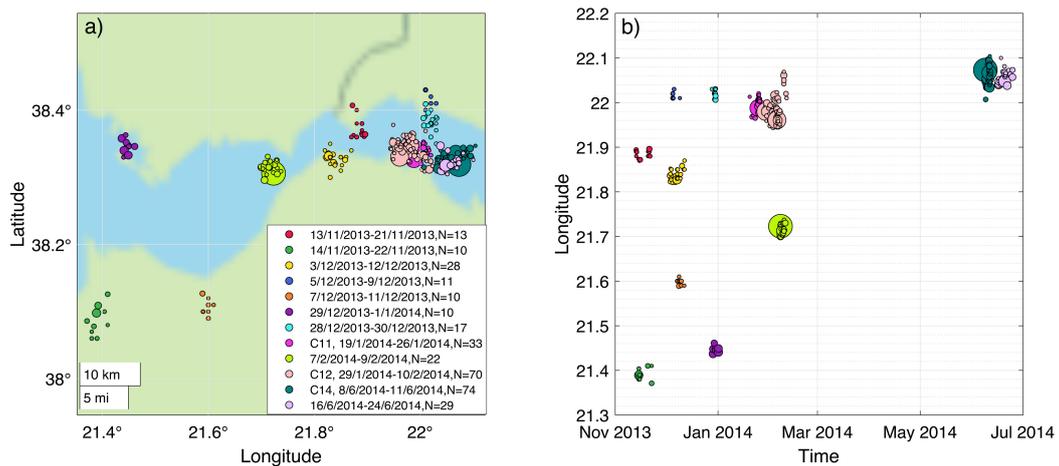


Figure C.3: a) Epicentral map of the seismic activity between November, 2013 and June, 2014. Twelve clusters with  $N \geq$  are occurred, including the  $C11$ ,  $C12$  and  $C14$  clusters. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

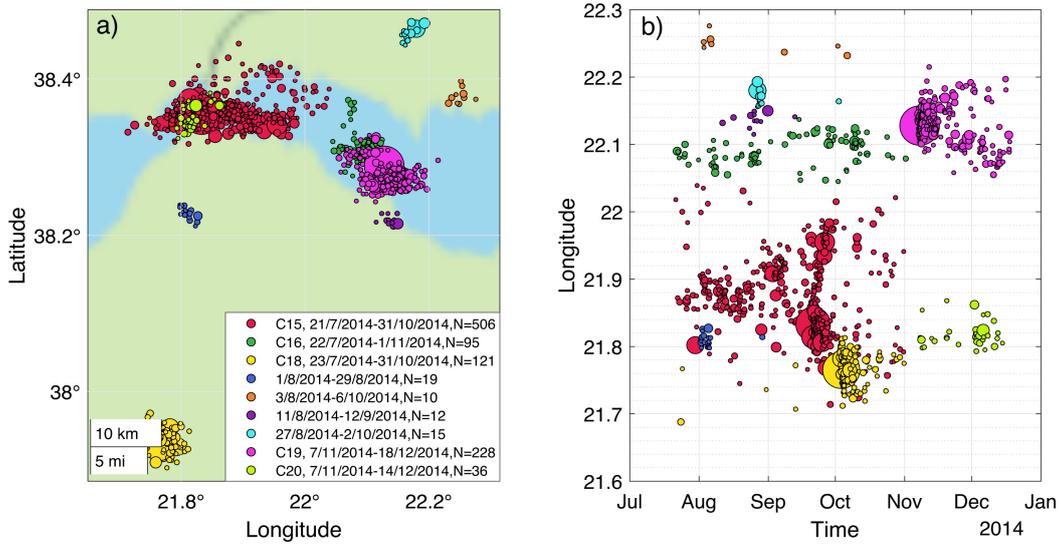


Figure C.4: a) Epicentral map of the intense seismic activity during the second half of 2014. Five major clusters are occurred, the *C15*, *C16*, *C18*, *C19* and *C20* and four smaller clusters with  $N \geq 10$  events. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

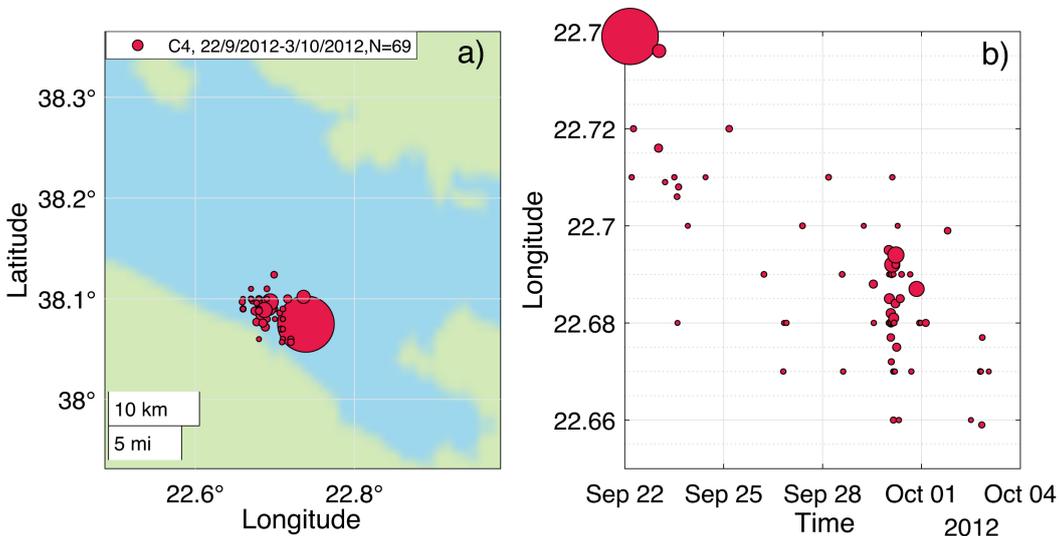


Figure C.5: a) Epicentral map of the seismic sequence Offshore Perachora. One major cluster, *C4*, including two sub sequences, the first initiated on 22 September and the second on 30 September 2012. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

C. APPENDIX C

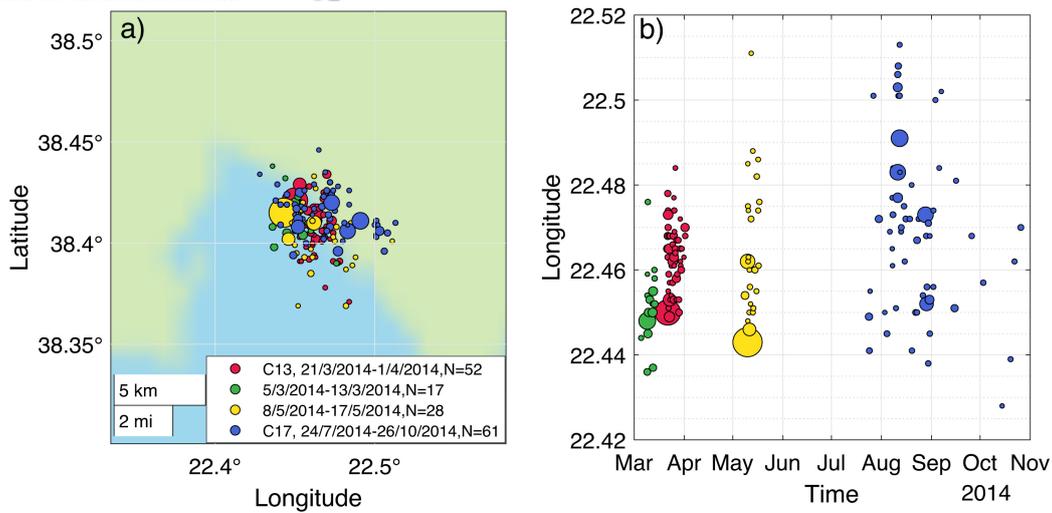


Figure C.6: a) Epicentral map of the seismic activity near Itea Gulf during 2014. Two major clusters are occurred, the *C13*, *C17* and four smaller ones with  $N \geq 10$  events. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

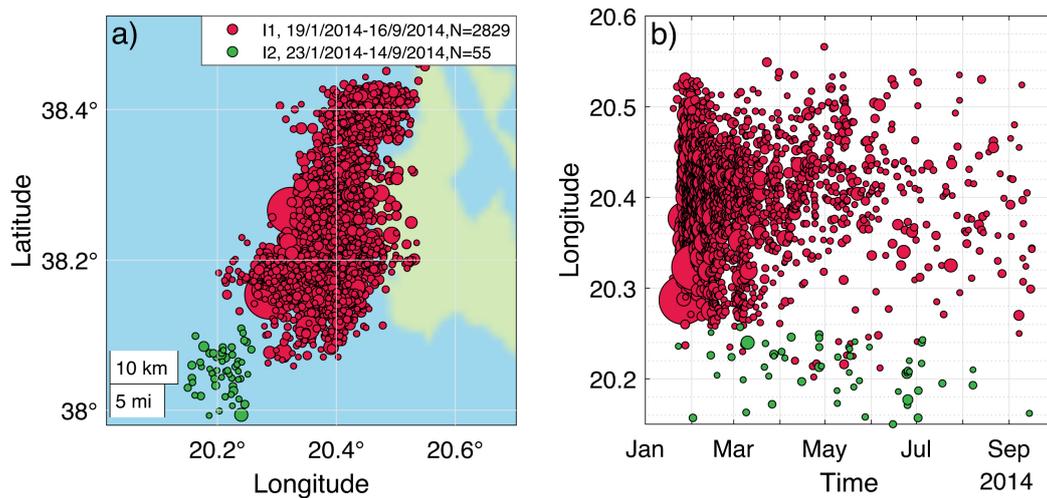


Figure C.7: a) Epicentral map of the 2014 Kefalonia earthquake sequence, *I1*, and a sub-cluster, *I2*, that occurred offshore the southern part of Kefalonia Island. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

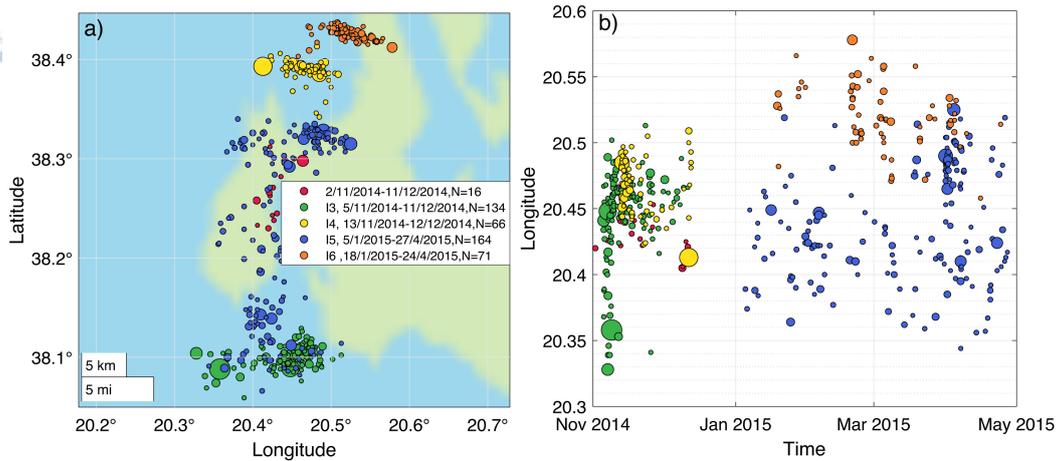


Figure C.8: a) Epicentral map of four main clusters,  $I_3$ ,  $I_4$ ,  $I_5$  and  $I_6$  with  $N \geq 30$  between November, 2014 and April, 2015. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

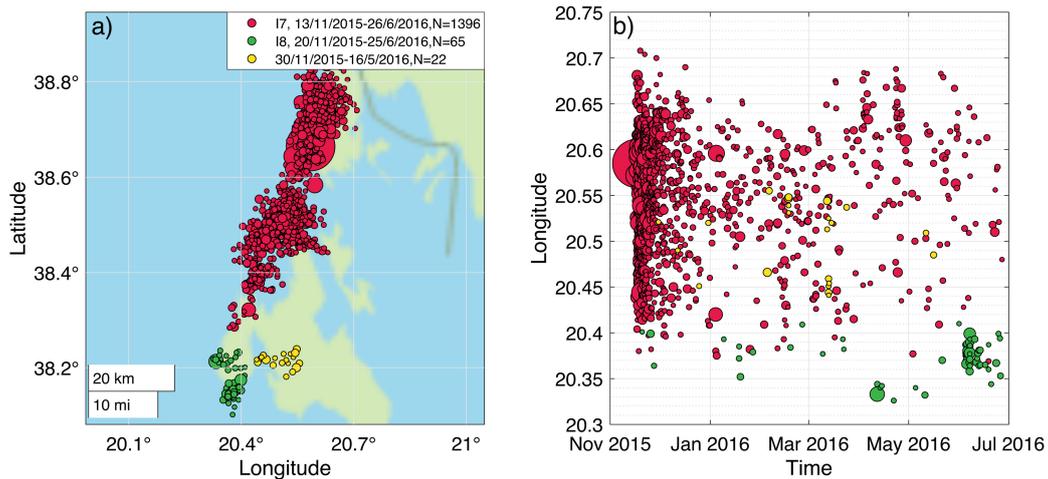


Figure C.9: a) Epicentral map of the 2015 Lefkada sequence,  $I_7$ , along with two sub-clusters in the southern part of Kefalonia Island. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

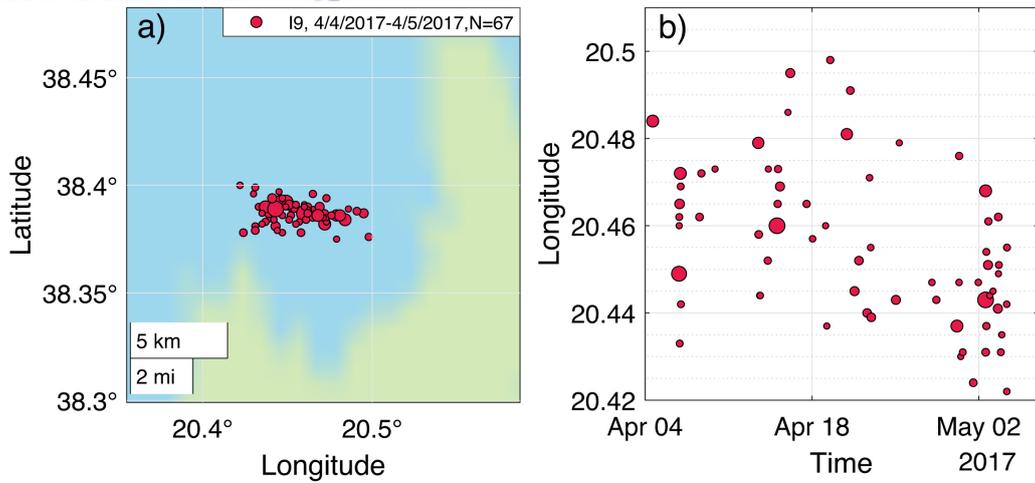


Figure C.10: a) Epicentral map of cluster *I9* located on the area between Lefkada and Kefalonia. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

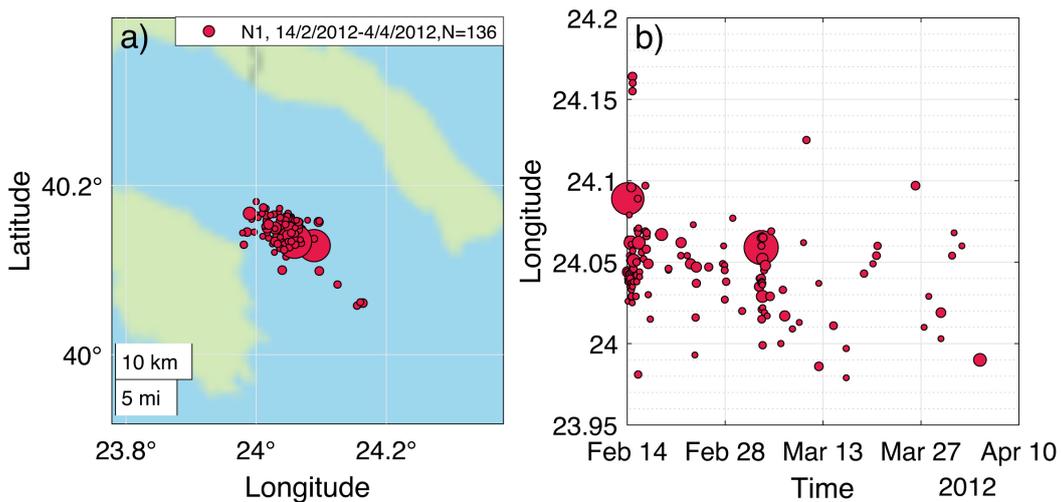


Figure C.11: a) Epicentral map of cluster *N1* comprised by two sub-sequences. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

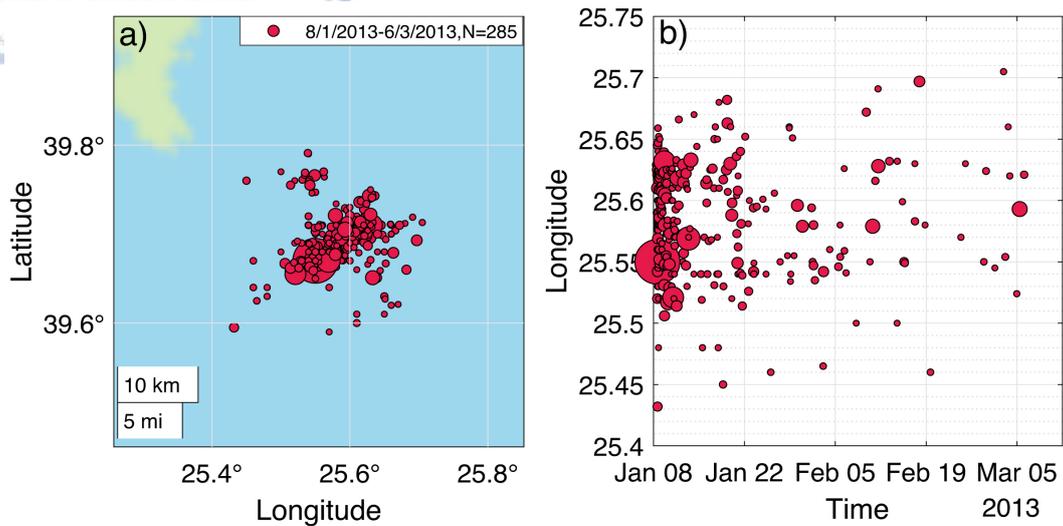


Figure C.12: a) Epicentral map of the 2013 North Aegean sequence, denoted *N3*. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

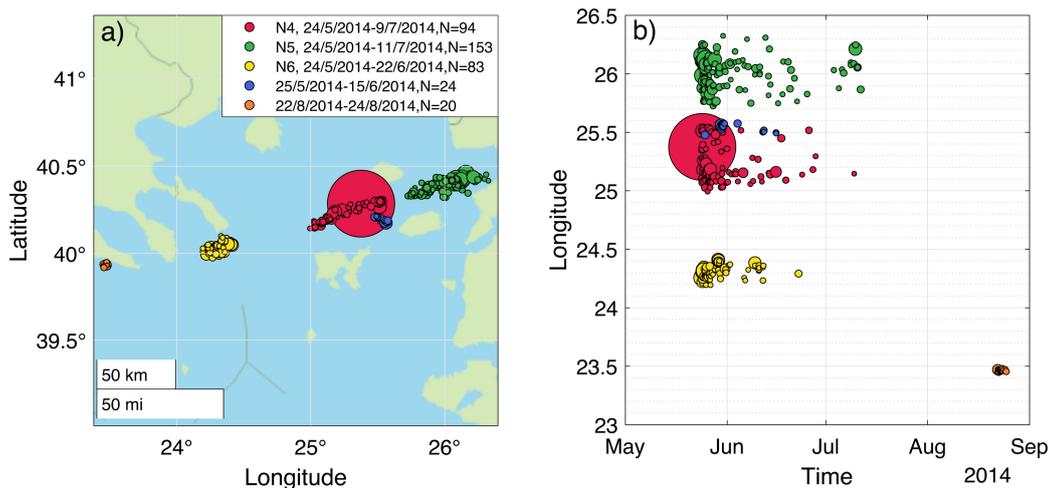


Figure C.13: a) Epicentral map of the 2013 North Aegean sequence, denoted *N3*. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

C. APPENDIX C

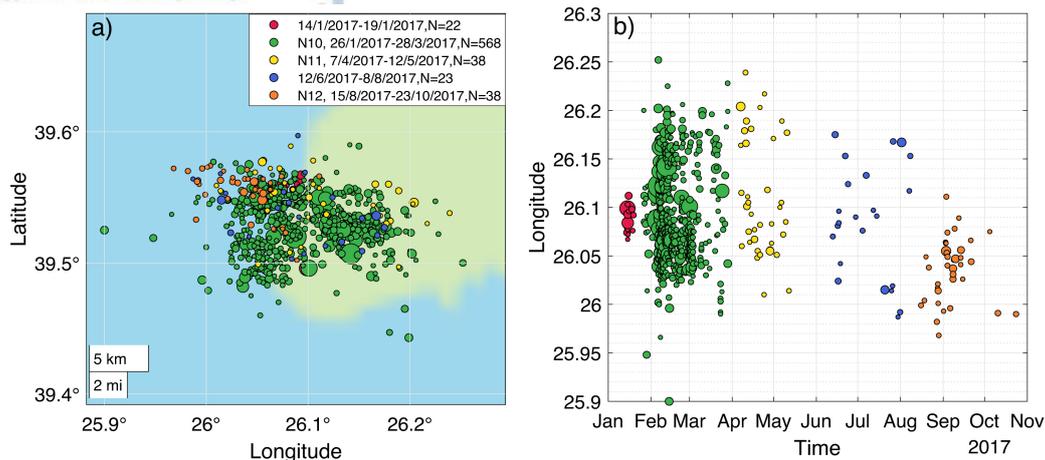


Figure C.14: a) Epicentral map of the seismic activity near the Aegean coast of NW Turkey during January–October 2017 confined into three clusters, *N10*, *N11* and *N12*. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

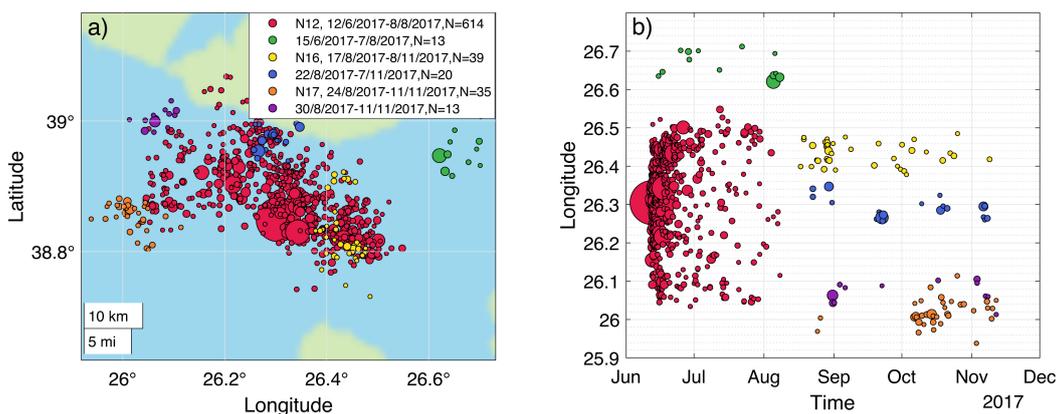


Figure C.15: a) Epicentral map of the seismic activity near the Aegean coast of NW Turkey during January–October 2017 confined into three clusters, *N10*, *N11* and *N12*. b) Space-time evolution of seismicity. Colours correspond to different clusters and the size of circles is proportional to the earthquakes magnitude.

## C.2 Additional fitting results in circum-Pacific belt

Table C.1: The log-likelihood and AIC values of the fitted MAP models to the events with  $M \geq 7.0, 7.6, 7.7, 7.8$  in circum-Pacific belt during 1918-2020.

# of S	$M_{thr}7.0$		# of S	$M_{thr} = 7.6$		$M_{thr}7.7$		$M_{thr} = 7.8$	
	LL	AIC		LL	AIC	LL	AIC	LL	AIC
3	1166	-2309.1	2	-3.8	19.6	-53.5	119.0	-73.5	159.1
4	1172	-2304.8	3	6.6	10.6	-45.4	114.9	-65.8	155.7

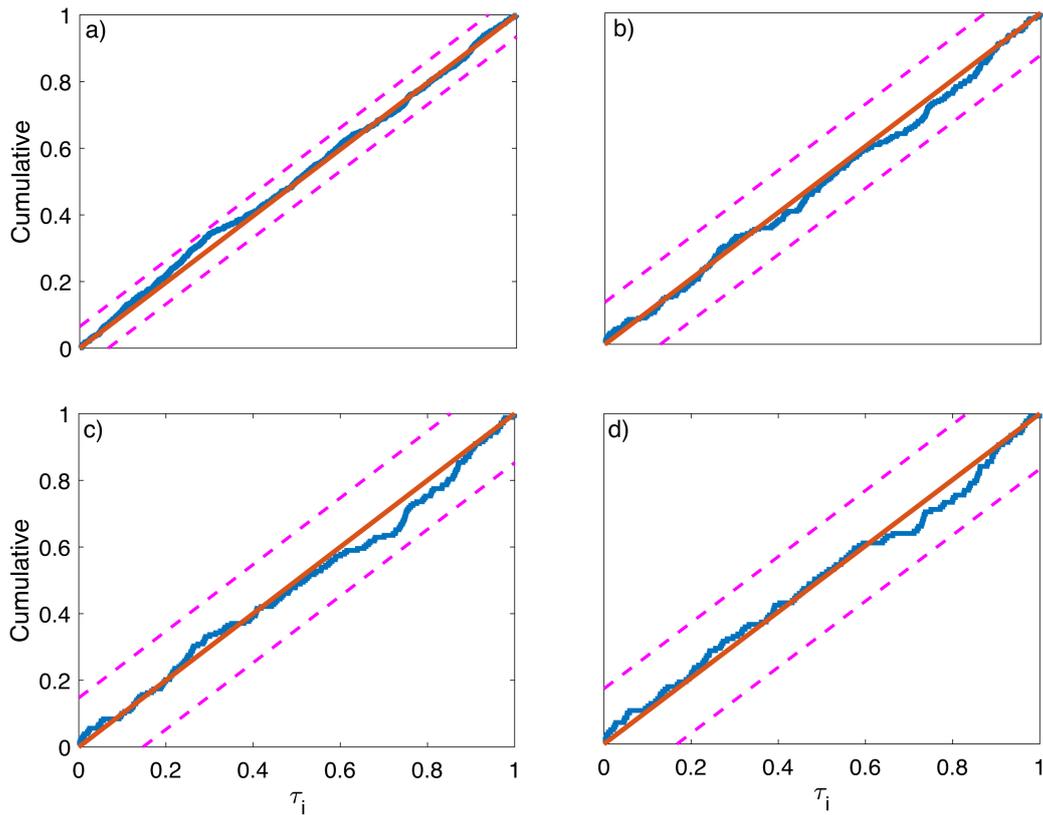


Figure C.16: Comparison between the cumulative number of residuals (blue step function) and the stationary Poisson process with unit rate (brown line) (a) for threshold  $M_{thr} = 7.0$ , (b)  $M_{thr} = 7.6$ , (c)  $M_{thr} = 7.7$  and (d)  $M_{thr} = 7.8$ . The pink dashed lines indicate the 95% confidence bounds.

C. APPENDIX C

Table C.2: P-values of the Runs and Kolmogorov-Smirnov tests implemented to the inter-event times of the residuals,  $E_i$ , for the three-states MAPs of the events with  $M \geq 7.0, 7.6, 7.7, 7.8$  in circum-Pacific belt during 1918-2020.

Statistical tests	$M_{thr} = 7.0$	$M_{thr} = 7.6$	$M_{thr} 7.7$	$M_{thr} = 7.8$
Runs test	0.442	0.804	0.819	0.947
KS test	0	$2.025 \cdot 10^{-4}$	0.003	0.054



## References

- Aden-Antoniów, F., Frank, W., and Seydoux, L. (2022). An adaptable random forest model for the declustering of earthquake catalogs. *Journal of Geophysical Research: Solid Earth*, 127(2):e2021JB023254.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.
- Aki, K. (1965). Maximum likelihood estimate of  $b$  in the formula  $\log N = a - bM$  and its confidence limits. *Bull. Earthq. Res. Inst. Tokyo Univ.*, 43:237–239.
- Ambraseys, N. and Jackson, J. (1997). Seismicity and strain in the Gulf of Corinth (Greece) since 1694. *Journal of Earthquake Engineering*, 1(3):433–474.
- Anderson, T. W. and Goodman, L. A. (1957). Statistical inference about Markov chains. *The Annals of Mathematical Statistics*, 28:89–110.
- Aristotle University of Thessaloniki (1981). Aristotle University of Thessaloniki Seismological Network. International Federation of Digital Seismograph Networks. dataset available at <http://dx.doi.org/doi:10.7914/SN/HT>.
- Artalejo, J. R., Gómez-Corral, A., and He, Q.-M. (2010). Markovian arrivals in stochastic modelling: A survey and some new results. *SORT*, 34(2):101–156.
- Asmussen, S. and Koole, G. (1993). Marked point processes as limits of Markovian arrival streams. *Journal of Applied Probability*, 30(2):365–372.

REFERENCES

---

- Asmussen, S., Nerman, O., and Olsson, M. (1996). Fitting phase-type distributions via the EM algorithm. *Scandinavian Journal of Statistics*, 23(4):419–441.
- Baiesi, M. and Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical review E*, 69(6):066106.
- Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The Annals of Mathematical Statistics*, 37(6):1554–1563.
- Bause, F. and Horváth, G. (2010). Fitting markovian arrival processes by incorporating correlation into phase type renewal processes. In *2010 Seventh International Conference on the Quantitative Evaluation of Systems*, pages 97–106. IEEE.
- Bayliss, K., Naylor, M., and Main, I. G. (2019). Probabilistic identification of earthquake clusters using rescaled nearest neighbour distance networks. *Geophysical Journal International*, 217(1):487–503.
- Bebbington, M. S. (2007). Identifying volcanic regimes using hidden Markov models. *Geophysical Journal International*, 171(2):921–942.
- Ben-Naim, E., Daub, E., and Johnson, P. (2013). Recurrence statistics of great earthquakes. *Geophysical Research Letters*, 40(12):3021–3025.
- Ben-Zion, Y. and Sammis, C. G. (2003). Characterization of fault zones. *Pure and applied geophysics*, 160(3):677–715.
- Ben-Zion, Y. and Zaliapin, I. (2020). Localization and coalescence of seismicity before large earthquakes. *Geophysical Journal International*, 223(1):561–583.
- Benali, A., Peresan, A., Varini, E., and Talbi, A. (2020). Modelling background seismicity components identified by nearest neighbour and stochastic declus-

REFERENCES

---

- tering approaches: The case of Northeastern Italy. *Stochastic Environmental Research and Risk Assessment*, 34(6):775–791.
- Bendick, R. and Bilham, R. (2017). Do weak global stresses synchronize earthquakes? *Geophysical Research Letters*, 44(16):8320–8327.
- Bendick, R. and Mencin, D. (2020). Evidence for synchronization in the global earthquake catalog. *Geophysical Research Letters*, 47(15):e2020GL087129.
- Bernard, P., Briole, P., Meyer, B., Lyon-Caen, H., Gomez, J.-M., Tiberi, C., Berge, C., Cattin, R., Hatzfeld, D., Lachet, C., et al. (1997). The Ms= 6.2, June 15, 1995 Aigion earthquake (Greece): evidence for low angle normal faulting in the Corinth rift. *Journal of Seismology*, 1(2):131–150.
- Beroza, G. C. (2012). How many great earthquakes should we expect? *Proceedings of the National Academy of Sciences*, 109(3):651–652.
- Bodrog, L., Buchholz, P., Heindl, A., Horváth, A., Horváth, G., Kolossváry, I., Mészáros, A., Németh, Z., Papp, J., Reinecke, P., Telek, M., and Vécsei, M. (2014). Program packages for computations with PH, ME distributions and MAP, RAP processes. User Manual. Available at <http://webspn.hit.bme.hu/~telek/tools/butools/>.
- Bodrog, L., Heindl, A., Horváth, G., and Telek, M. (2008). A Markovian canonical form of second-order matrix-exponential processes. *European Journal of Operational Research*, 190(2):459–477.
- Bottiglieri, M., Lippiello, E., Godano, C., and De Arcangelis, L. (2009). Identification and spatiotemporal organization of aftershocks. *Journal of Geophysical Research: Solid Earth*, 114(B3).
- Bountzis, P., Kostoglou, A., Papadimitriou, E., and Karakostas, V. (2021). Identification of spatiotemporal seismicity clusters in central Ionian Islands (Greece). *Physics of the Earth and Planetary Interiors*, 312:106675.

REFERENCES

---

- Bountzlis, P., Papadimitriou, E., and Tsaklidis, G. (2019). Estimating the earthquake occurrence rates in Corinth Gulf (Greece) through Markovian arrival process modeling. *Journal of Applied Statistics*, 46(6):995–1020.
- Bourouis, S. and Cornet, F. H. (2009). Microseismic activity and fluid fault interactions: some results from the Corinth Rift Laboratory (CRL), Greece. *Geophysical Journal International*, 178(1):561–580.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Bradley, J. V. (1968). Distribution-free statistical tests.
- Bražėnas, M., Horváth, G., and Telek, M. (2018). Parallel algorithms for fitting Markov arrival processes. *Performance Evaluation*, 123:50–67.
- Breuer, L. and Kume, A. (2010). An EM algorithm for Markovian arrival processes observed at discrete times. In *International GI/ITG Conference on Measurement, Modelling, and Evaluation of Computing Systems and Dependability and Fault Tolerance*, pages 242–258. Springer.
- Briole, P., Rigo, A., Lyon-Caen, H., Ruegg, J. C., Papazissi, K., Mitsakaki, C., Balodimou, A., Veis, G., Hatzfeld, D., and Deschamps, A. (2000). Active deformation of the Corinth rift, Greece: results from repeated Global Positioning System surveys between 1990 and 1995. *Journal of Geophysical Research: Solid Earth*, 105(B11):25605–25625.
- Buchholz, P. (2003). An EM-algorithm for MAP fitting from real traffic data. In *International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, pages 218–236. Springer.
- Buchholz, P. and Kriege, J. (2009). A heuristic approach for fitting MAPs to moments and joint moments. In *2009 Sixth international conference on the quantitative evaluation of systems*, pages 53–62. IEEE.

REFERENCES

---

- Buchholz, P., Kriege, J., and Felko, I. (2014). *Input Modeling with Phase-Type Distributions and Markov Models: Theory and Applications*. Springer, Cham.
- Bufe, C. G. and Perkins, D. M. (2005). Evidence for a global seismic-moment release sequence. *Bulletin of the Seismological Society of America*, 95(3):833–843.
- Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference*. Springer-Verlag, New York, 2nd edition.
- Casale, G., Zhang, E. Z., and Smirni, E. (2010). Trace data characterization and fitting for Markov modeling. *Performance Evaluation*, 67(2):61–79.
- Cesca, S. (2020). Seiscloud, a tool for density-based seismicity clustering and visualization. *Journal of Seismology*, 24(3):443–457.
- Cesca, S., Grigoli, F., Heimann, S., Dahm, T., Kriegerowski, M., Sobiesiak, M., Tassara, C., and Olcay, M. (2016). The Mw 8.1 2014 Iquique, Chile, seismic sequence: a tale of foreshocks and aftershocks. *Geophysical Journal International*, 204(3):1766–1780.
- Cesca, S., Şen, A. T., and Dahm, T. (2014). Seismicity monitoring by cluster analysis of moment tensors. *Geophysical Journal International*, 196(3):1813–1826.
- Chakravarti, I., Laha, R., and Roy, J. (1967). *Handbook of Methods of Applied Statistics-Volume 1: Techniques of Computation, Descriptive Methods, and Statistical Inference*. John Wiley & Sons, Inc., New York-London-Sydney.
- Chambers, D. W., Baglivo, J. A., Ebel, J. E., and Kafka, A. L. (2012). Earthquake forecasting using hidden Markov models. *Pure and applied geophysics*, 169(4):625–639.

REFERENCES

---

- Chen, Y., Liu, M., and Luo, G. (2020). Complex temporal patterns of large earthquakes: Devil's staircases. *Bulletin of the Seismological Society of America*, 110(3):1064–1076.
- Chousianitis, K., Ganas, A., and Evangelidis, C. P. (2015). Strain and rotation rate patterns of mainland Greece from continuous GPS data and comparison between seismic and geodetic moment release. *Journal of Geophysical Research: Solid Earth*, 120(5):3909–3931.
- Console, R., Carluccio, R., Papadimitriou, E., and Karakostas, V. (2015). Synthetic earthquake catalogs simulating seismic activity in the Corinth Gulf, Greece, fault system. *Journal of Geophysical Research: Solid Earth*, 120(1):326–343.
- Console, R., Falcone, G., Karakostas, V., Murru, M., Papadimitriou, E., and Rhoades, D. (2013). Renewal models and coseismic stress transfer in the Corinth Gulf, Greece, fault system. *Journal of Geophysical Research: Solid Earth*, 118(7):3655–3673.
- Console, R., Jackson, D., and Kagan, Y. (2010). Using the ETAS model for catalog declustering and seismic background assessment. *Pure and applied geophysics*, 167(6):819–830.
- Cordeiro, J. D. and Kharoufeh, J. P. (2011). Batch Markovian Arrival Processes (BMAP). *Wiley Encyclopedia of Operations Research and Management Science*.
- Crespo Martín, C. and Martín-González, F. (2021). Statistical Analysis of Intraplate Seismic Clusters: The Case of the NW Iberian Peninsula. *Pure and Applied Geophysics*, 178(9):3355–3374.
- Daley, D. J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer.

REFERENCES

---

- Daley, D. J. and Vere-Jones, D. (2004). Scoring probability forecasts for point processes: The entropy score and information gain. *Journal of Applied Probability*, 41(A):297–312.
- Dascher-Cousineau, K., Brodsky, E. E., Lay, T., and Goebel, T. H. (2020). What controls variations in aftershock productivity? *Journal of Geophysical Research: Solid Earth*, 125(2):e2019JB018111.
- Daub, E. G., Ben-Naim, E., Guyer, R. A., and Johnson, P. A. (2012). Are megaquakes clustered? *Geophysical Research Letters*, 39(6):L06308.
- Daub, E. G., Trugman, D. T., and Johnson, P. A. (2015). Statistical tests on clustered global earthquake synthetic data sets. *Journal of Geophysical Research: Solid Earth*, 120(8):5693–5716.
- Davoudi, N., Tavakoli, H. R., Zare, M., and Jalilian, A. (2018). Declustering of Iran earthquake catalog (1983–2017) using the epidemic-type aftershock sequence (ETAS) model. *Acta Geophysica*, 66(6):1359–1373.
- De Barros, L., Cappa, F., Deschamps, A., and Dublanchet, P. (2020). Imbricated aseismic slip and fluid diffusion drive a seismic swarm in the Corinth Gulf, Greece. *Geophysical Research Letters*, 47(9):e2020GL087142.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B Statistical Methodology*, 39(1):1–38.
- Di Giacomo, D., Engdahl, E. R., and Storchak, D. A. (2018). The ISC-GEM earthquake catalogue (1904–2014): status after the extension project. *Earth System Science Data*, 10(4):1877–1899.
- Dimer de Oliveira, F. (2012). Can we trust earthquake cluster detection tests? *Geophysical research letters*, 39(17).

REFERENCES

---

Duverger, C., Lambotte, S., Bernard, P., Lyon-Caen, H., Deschamps, A., and Necessian, A. (2018). Dynamics of microseismicity and its relationship with the active structures in the western Corinth Rift (Greece). *Geophysical Journal International*, 215(1):196–221.

Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 58(1):25–45.

Ebel, J. E., Chambers, D. W., Kafka, A. L., and Baglivo, J. A. (2007). Non-Poissonian earthquake clustering and the hidden Markov model as bases for earthquake forecasting in California. *Seismological Research Letters*, 78(1):57–65.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, volume 96, pages 226–231. AAAI Press.

Evangelidis, C. (2015). Imaging supershear rupture for the 2014 Mw6.9 Northern Aegean earthquake by backprojection of strong motion waveforms. *Geophysical Research Letters*, 42(2):307–315.

Faenza, L., Marzocchi, W., and Boschi, E. (2003). A non-parametric hazard model to characterize the spatio-temporal occurrence of large earthquakes; an application to the italian catalogue. *Geophysical Journal International*, 155(2):521–531.

Faenza, L., Marzocchi, W., Lombardi, A. M., and Console, R. (2004). Some insights into the time clustering of large earthquakes in italy. *Annals of Geophysics*, 47(5):1635–1640.

Faenza, L., Marzocchi, W., Serretti, P., and Boschi, E. (2008). On the spatio-

REFERENCES

---

- temporal distribution of M7.0+ worldwide seismicity with a non-parametric statistics. *Tectonophysics*, 449(1-4):97–104.
- Feldmann, A. and Whitt, W. (1998). Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance evaluation*, 31(3-4):245–279.
- Felzer, K. R. and Brodsky, E. E. (2006). Decay of aftershock density with distance indicates triggering by dynamic stress. *Nature*, 441(7094):735–738.
- Fiedler, B., Zöller, G., Holschneider, M., and Hainzl, S. (2018). Multiple change-point detection in spatiotemporal seismicity data. *Bulletin of the Seismological Society of America*, 108(3A):1147–1159.
- Field, E. H., Milner, K. R., Hardebeck, J. L., Page, M. T., van der Elst, N., Jordan, T. H., Michael, A. J., Shaw, B. E., and Werner, M. J. (2017). A spatiotemporal clustering model for the third Uniform California Earthquake Rupture Forecast (UCERF3-ETAS): Toward an operational earthquake forecast. *Bulletin of the Seismological Society of America*, 107(3):1049–1081.
- Finch, S. J., Mendell, N. R., and Thode Jr, H. C. (1989). Probabilistic measures of adequacy of a numerical search for a global maximum. *Journal of the American Statistical Association*, 84(408):1020–1023.
- Fischer, W. and Meier-Hellstern, K. (1993). The Markov-modulated Poisson process (MMPP) cookbook. *Performance evaluation*, 18(2):149–171.
- Fortunato, S. and Hric, D. (2016). Community detection in networks: A user guide. *Physics reports*, 659:1–44.
- Frohlich, C. and Davis, S. D. (1990). Single-link cluster analysis as a method to evaluate spatial and temporal properties of earthquake catalogues. *Geophysical Journal International*, 100(1):19–32.

## REFERENCES

---

- Ganas, A., Chousianitis, K., Batsi, E., Kolligri, M., Agalos, A., Chouliaras, G., and Makropoulos, K. (2013). The January 2010 Efpalion earthquakes (Gulf of Corinth, Central Greece): earthquake interactions and blind normal faulting. *Journal of seismology*, 17(2):465–484.
- Ganas, A., Serpelloni, E., Drakatos, G., Kolligri, M., Adamis, I., Tsimi, C., and Batsi, E. (2009). The Mw 6.4 SW-Achaia (western Greece) earthquake of 8 June 2008: Seismological, field, GPS observations, and stress modeling. *Journal of Earthquake Engineering*, 13(8):1101–1124.
- Gardner, J. and 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.
- Gentili, S. and Bressan, G. (2008). The partitioning of radiated energy and the largest aftershock of seismic sequences occurred in the northeastern Italy and western Slovenia. *Journal of Seismology*, 12(3):343–354.
- Georgoulas, G., Konstantaras, A., Katsifarakis, E., Stylios, C. D., Maravelakis, E., and Vachtsevanos, G. J. (2013). “seismic-mass” density-based algorithm for spatio-temporal clustering. *Expert Systems with Applications*, 40(10):4183–4189.
- Gerstenberger, M. C., Marzocchi, W., Allen, T., Pagani, M., Adams, J., Danciu, L., Field, E. H., Fujiwara, H., Luco, N., Ma, K.-F., et al. (2020). Probabilistic seismic hazard analysis at regional and national scales: State of the art and future challenges. *Reviews of Geophysics*, 58(2):e2019RG000653.
- Gerstenberger, M. C., Wiemer, S., Jones, L. M., and Reasenber, P. A. (2005). Real-time forecasts of tomorrow’s earthquakes in California. *Nature*, 435(7040):328–331.
- Godano, C. and Tramelli, A. (2016). How long is an aftershock sequence? *Pure and Applied Geophysics*, 173(7):2295–2304.

REFERENCES

---

- Gomberg, J. and Sherrod, B. (2014). Crustal earthquake triggering by modern great earthquakes on subduction zone thrusts. *Journal of Geophysical Research: Solid Earth*, 119(2):1235–1250.
- Grund, M., Groos, J. C., and Ritter, J. R. (2016). Fault reactivation analysis using microearthquake clustering based on signal-to-noise weighted waveform similarity. *Pure and applied geophysics*, 173(7):2325–2355.
- Gupta, A. and Baker, J. W. (2017). Estimating spatially varying event rates with a change point using bayesian statistics: Application to induced seismicity. *Structural safety*, 65:1–11.
- Gutenberg, R. and Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34:185–188.
- Hainzl, S. (2004). Seismicity patterns of earthquake swarms due to fluid intrusion and stress triggering. *Geophysical Journal International*, 159(3):1090–1096.
- Hainzl, S. and Ogata, Y. (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling. *Journal of Geophysical Research: Solid Earth*, 110(B5):B05S07.
- Hainzl, S., Zakharova, O., and Marsan, D. (2013). Impact of aseismic transients on the estimation of aftershock productivity parameters. *Bulletin of the Seismological Society of America*, 103(3):1723–1732.
- Hanks, T. C. and Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research: Solid Earth*, 84(B5):2348–2350.
- Hardebeck, J. L., Llenos, A. L., Michael, A. J., Page, M. T., and Van Der Elst, N. (2019). Updated California aftershock parameters. *Seismological Research Letters*, 90(1):262–270.

REFERENCES

---

Harte, D. (2017). HiddenMarkov: Hidden Markov Models. *R package version*, 1(7).

Hatzfeld, D., Karakostas, V., Ziazia, M., Kassaras, I., Papadimitriou, E., Makropoulos, K., Voulgaris, N., and Papaioannou, C. (2000). Microseismicity and faulting geometry in the Gulf of Corinth (Greece). *Geophysical Journal International*, 141(2):438–456.

Hatzfeld, D., Kementzetzidou, D., Karakostas, V., Ziazia, M., Nothard, S., Diagourtas, D., Deschamps, A., Karakaisis, G., Papadimitriou, P., Scordilis, M., et al. (1996). The Galaxidi earthquake of 18 November 1992: A possible asperity within the normal fault system of the Gulf of Corinth (Greece). *B. Seismol. Soc. Am.*, 86(6):1987–1991.

Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 33(3):438–443.

He, Q.-M. (2010). Construction of continuous time Markovian arrival processes. *Journal of Systems Science and Systems Engineering*, 19(3):351–366.

Heindl, A. (2004). Inverse characterization of hyperexponential MAP(2)s. In *Proc. 11th Int. Conference on Analytical and Stochastic Modelling Techniques and Applications (ASMTA)*, volume 31. Citeseer.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3:1163–1174.

Hill, D. and Prejean, S. (2007). Dynamic triggering in treatise of geophysics, vol. 4. ed. *Kanamori, H., Elsevier Science*, pages 257–291.

Horváth, G., Buchholz, P., and Telek, M. (2005). A MAP fitting approach with independent approximation of the inter-arrival time distribution and the lag correlation. In *Second International Conference on the Quantitative Evaluation of Systems (QEST'05)*, pages 124–133. IEEE.

## REFERENCES

---

- Horváth, G. and Okamura, H. (2013). A fast em algorithm for fitting marked markovian arrival processes with a new special structure. In *Comput. Perform. Eng. EPEW*, volume 8168 of *Lecture Notes in Computer Science*, pages 119–133, Berlin, Heidelberg. Springer.
- Huang, Q., Gerstenberger, M., and Zhuang, J. (2016). Current challenges in statistical seismology. *Pure and Applied Geophysics*, 173(1):1–3.
- Iliopoulos, A., Chorozoglou, D., Kourouklas, C., Mangira, O., and Papadimitriou, E. (2020). Memory and renewal aging of strong earthquakes in Hellenic seismicity. *Chaos, Solitons & Fractals*, 131:109511.
- Jackson, J., Gagnepain, J., Houseman, G., King, G., Papadimitriou, P., Soufleris, C., and Virieux, J. (1982). Seismicity, normal faulting, and the geomorphological development of the Gulf of Corinth (Greece): the Corinth earthquakes of February and March 1981. *Earth and Planetary Science Letters*, 57(2):377–397.
- Jackson, J. and McKenzie, D. (1988). The relationship between plate motions and seismic moment tensors, and the rates of active deformation in the Mediterranean and Middle East. *Geophysical Journal International*, 93(1):45–73.
- Jacobs, K. M., Smith, E. G., Savage, M. K., and Zhuang, J. (2013). Cumulative rate analysis (CURATE): A clustering algorithm for swarm dominated catalogs. *Journal of Geophysical Research: Solid Earth*, 118(2):553–569.
- Jordan, T., Chen, Y.-T., Gasparini, P., Madariaga, R., Main, I., Marzocchi, W., Papadopoulos, G., Yamaoka, K., Zschau, J., et al. (2011). Operational earthquake forecasting: State of knowledge and guidelines for implementation. *Annals of Geophysics*, 54(4).
- Kadane, J. B. and Lazar, N. A. (2004). Methods and criteria for model selection. *Journal of the American statistical Association*, 99(465):279–290.

## REFERENCES

---

- Kagan, Y. Y. and Jackson, D. D. (1991). Long-term earthquake clustering. *Geophysical Journal International*, 104(1):117–133.
- Kagan, Y. Y. and Knopoff, L. (1981). Stochastic synthesis of earthquake catalogs. *Journal of Geophysical Research: Solid Earth*, 86(B4):2853–2862.
- Kahle, H.-G., Müller, M. V., Geiger, A., Danuser, G., Mueller, S., Veis, G., Billiris, H., and Paradissis, D. (1995). The strain field in northwestern Greece and the Ionian Islands: results inferred from GPS measurements. *Tectonophysics*, 249(1-2):41–52.
- Kamer, Y., Ouillon, G., and Sornette, D. (2020). Fault network reconstruction using agglomerative clustering: applications to southern Californian seismicity. *Natural Hazards and Earth System Sciences*, 20(12):3611–3625.
- Kang, S. H., Kim, Y. H., Sung, D. K., and Choi, B. D. (2002). An application of Markovian arrival process (MAP) to modeling superposed ATM cell streams. *IEEE Transactions on Communications*, 50(4):633–642.
- Kapetanidis, V. (2017). *Spatiotemporal patterns of microseismicity for the identification of active fault structures using seismic waveform cross-correlation and double-difference relocation*. PhD thesis.
- Kapetanidis, V., Deschamps, A., Papadimitriou, P., Matrullo, E., Karakonstantis, A., Bozionelos, G., Kaviris, G., Serpetsidaki, A., Lyon-Caen, H., Voulgaris, N., et al. (2015). The 2013 earthquake swarm in Helike, Greece: seismic activity at the root of old normal faults. *Geophysical Journal International*, 202(3):2044–2073.
- Kapetanidis, V., Michas, G., Kaviris, G., and Vallianatos, F. (2021). Spatiotemporal Properties of Seismicity and Variations of Shear-Wave Splitting Parameters in the Western Gulf of Corinth (Greece). *Applied Sciences*, 11(14):6573.

REFERENCES

---

Karakostas, B., Papadimitriou, E., Hatzfeld, D., Makaris, D., Makropoulos, K., Diagourtas, D., Papaioannou, C., Stavrakakis, G., Drakopoulos, J., and Papazachos, B. (1994). The aftershock sequence and focal properties of the July 14, 1993 ( $M_s=5.4$ ) Patras earthquake. In *Proceedings of the 7th Congress, Thessaloniki, May 1994*, Bulletin of the Geological Society of Greece, pages 167–174.

Karakostas, V., Karagianni, E., and Paradisopoulou, P. (2012). Space-time analysis, faulting and triggering of the 2010 earthquake doublet in western Corinth Gulf. *Natural Hazards*, 63(2):1181–1202.

Karakostas, V., Mirek, K., Mesimeri, M., Papadimitriou, E., and Mirek, J. (2017). The aftershock sequence of the 2008 Achaia, Greece, earthquake: joint analysis of seismicity relocation and persistent scatterers interferometry. *Pure and Applied Geophysics*, 174(1):151–176.

Karakostas, V., Papadimitriou, E., and Gospodinov, D. (2014). Modelling the 2013 North Aegean (Greece) seismic sequence: geometrical and frictional constraints, and aftershock probabilities. *Geophysical Journal International*, 197(1):525–541.

Karakostas, V., Papadimitriou, E., Mesimeri, M., Gkarlaouni, C., and Paradisopoulou, P. (2015). The 2014 Kefalonia doublet ( $M_w$  6.1 and  $M_w$  6.0), central Ionian Islands, Greece: Seismotectonic implications along the Kefalonia transform fault zone. *Acta Geophysica*, 63(1):1–16.

Karakostas, V. G. and Papadimitriou, E. E. (2010). Fault complexity associated with the 14 August 2003  $M_w$ 6.2 Lefkada, Greece, aftershock sequence. *Acta geophysica*, 58(5):838–854.

Kato, A. and Ben-Zion, Y. (2021). The generation of large earthquakes. *Nature Reviews Earth & Environment*, 2(1):26–39.

REFERENCES

---

- Kaviris, G., Millas, C., Spingos, I., Kapetanidis, V., Fountoulakis, I., Papadimitriou, P., Voulgaris, N., and Makropoulos, K. (2018). Observations of shear-wave splitting parameters in the Western Gulf of Corinth focusing on the 2014 Mw=5.0 earthquake. *Physics of the Earth and Planetary Interiors*, 282:60–76.
- Kerr, R. A. (2011). More megaquakes on the way? That depends on your statistics. *Science*, 332:411.
- Kiratzi, A., Sokos, E., Ganas, A., Tselentis, A., Benetatos, C., Roumelioti, Z., Serpetsidaki, A., Andriopoulos, G., Galanis, O., and Petrou, P. (2008). The April 2007 earthquake swarm near Lake Trichonis and implications for active tectonics in western Greece. *Tectonophysics*, 452(1-4):51–65.
- Kiratzi, A., Tsakiroudi, E., Benetatos, C., and Karakaisis, G. (2016). The 24 May 2014 (Mw6. 8) earthquake (North Aegean Trough): spatiotemporal evolution, source and slip model from teleseismic data. *Physics and Chemistry of the Earth, Parts A/B/C*, 95:85–100.
- Kiratzi, A. A. and Langston, C. A. (1991). Moment tensor inversion of the 1983 January 17 Kefallinia event of Ionian islands (Greece). *Geophysical Journal International*, 105(2):529–535.
- Kostoglou, A., Bountzis, P., Karakostas, V., and Papadimitriou, E. (2021). Classification of earthquake repeaters in the central Ionian Islands area. In *EGU General Assembly Conference Abstracts*, pages EGU21–11966.
- Kostoglou, A., Bountzis, P., Karakostas, V. G., and Papadimitriou, E. E. (2020). Microseismicity Study Of The Central Ionian Islands (Greece) Accompanied By A Focal Mechanism Clustering Investigation. In *AGU Fall Meeting Abstracts*, volume 2020, pages S047–0001.
- Kumazawa, T. and Ogata, Y. (2014). Nonstationary ETAS models for non-standard earthquakes. *The Annals of Applied Statistics*, 8(3):1825–1852.

REFERENCES

---

- Lay, T. (2015). The surge of great earthquakes from 2004 to 2014. *Earth and Planetary Science Letters*, 409:133–146.
- Le Pichon, X. and Angelier, J. (1979). The Hellenic arc and trench system: a key to the neotectonic evolution of the eastern Mediterranean area. *Tectonophysics*, 60(1-2):1–42.
- Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64(325):387–389.
- Lippiello, E. and Bountzlis, P. (2021). An objective criterion for cluster detection in stochastic epidemic models. *arXiv preprint arXiv:2104.04138*.
- Lippiello, E., Giacco, F., Arcangelis, L. d., Marzocchi, W., and Godano, C. (2014). Parameter estimation in the etas model: Approximations and novel methods. *Bulletin of the Seismological Society of America*, 104(2):985–994.
- Lippiello, E., Godano, C., and de Arcangelis, L. (2019). The relevance of foreshocks in earthquake triggering: A statistical study. *Entropy*, 21(2).
- Lippiello, E., Marzocchi, W., De Arcangelis, L., and Godano, C. (2012). Spatial organization of foreshocks as a tool to forecast large earthquakes. *Scientific reports*, 2(1):1–6.
- Llenos, A. L., McGuire, J. J., and Ogata, Y. (2009). Modeling seismic swarms triggered by aseismic transients. *Earth and Planetary Science Letters*, 281(1-2):59–69.
- Llenos, A. L. and Michael, A. J. (2017). Forecasting the (un) productivity of the 2014 M6.0 South Napa aftershock sequence. *Seismological Research Letters*, 88(5):1241–1251.
- Llenos, A. L. and Michael, A. J. (2020). Regionally optimized background earthquake rates from ETAS (ROBERE) for probabilistic seismic hazard

REFERENCES

---

assessment. *Bulletin of the Seismological Society of America*, 110(3):1172–1190.

Lombardi, A. M. and Marzocchi, W. (2007). Evidence of clustering and non-stationarity in the time distribution of large worldwide earthquakes. *Journal of Geophysical Research: Solid Earth*, 112:B02303.

Lombardi, A. M. and Marzocchi, W. (2010). The ETAS model for daily forecasting of Italian seismicity in the CSEP experiment. *Annals of Geophysics*, 53(3):155–164.

Louvari, E., Kiratzi, A., and Papazachos, B. (1999). The Cephalonia transform fault and its extension to western Lefkada Island (Greece). *Tectonophysics*, 308(1-2):223–236.

Lu, S. (2012). Markov modulated Poisson process associated with state-dependent marks and its applications to the deep earthquakes. *Annals of the Institute of Statistical Mathematics*, 64(1):87–106.

Lu, S. (2017a). A continuous-time HMM approach to modeling the magnitude-frequency distribution of earthquakes. *Journal of Applied Statistics*, 44(1):71–88.

Lu, S. (2017b). Long-term b value variations of shallow earthquakes in New Zealand: a HMM-based analysis. *Pure and Applied Geophysics*, 174(4):1629–1641.

Lu, S. (2019). A bayesian multiple changepoint model for marked poisson processes with applications to deep earthquakes. *Stochastic Environmental Research and Risk Assessment*, 33(1):59–72.

Lu, S. and Vere-Jones, D. (2011). Large occurrence patterns of New Zealand deep earthquakes: Characterization by use of a switching Poisson model. *Pure App. Geophys.*, 168(10):1567–1585.

REFERENCES

---

Lucantoni, D. M. (1991). New results on the single server queue with a batch Markovian arrival process. *Communications in Statistics. Stochastic Models*, 7(1):1–46.

Lucantoni, D. M., Meier-Hellstern, K. S., and Neuts, M. F. (1990). A single-server queue with server vacations and a class of non-renewal arrival processes. *Advances in Applied Probability*, 22(3):676–705.

Luginbuhl, M., Rundle, J. B., Hawkins, A., and Turcotte, D. L. (2018). Nowcasting earthquakes: a comparison of induced earthquakes in Oklahoma and at the Geysers, California. *Pure and Applied Geophysics*, 175(1):49–65.

Lykou, R., Tsaklidis, G., and Papadimitriou, E. (2020). Change point analysis on the Corinth Gulf (Greece) seismicity. *Physica A: Statistical Mechanics and its Applications*, 541:123630.

L'vov, V. S., Pomyalov, A., and Procaccia, I. (2001). Outliers, extreme events, and multiscaling. *Physical Review E*, 63(5):056118.

MacDonald, I. L. and Zucchini, W. (1997). *Hidden Markov and other models for discrete-valued time series*. Chapman & Hall, London, U.K.

Marsan, D. and Lengline, O. (2008). Extending earthquakes' reach through cascading. *Science*, 319(5866):1076–1079.

Marsan, D., Reverso, T., Helmstetter, A., and Enescu, B. (2013). Slow slip and aseismic deformation episodes associated with the subducting pacific plate offshore japan, revealed by changes in seismicity. *Journal of Geophysical Research: Solid Earth*, 118(9):4900–4909.

Martínez-Garzón, P., Ben-Zion, Y., Zaliapin, I., and Bohnhoff, M. (2019). Seismic clustering in the sea of marmara: Implications for monitoring earthquake processes. *Tectonophysics*, 768:228176.

REFERENCES

---

- McKenzie, D. (1972). Active tectonics of the mediterranean region. *Geophysical Journal International*, 30(2):109–185.
- Mendoza-Rosas, A., la Cruz-Reyna, D., et al. (2009). A mixture of exponentials distribution for a simple and precise assessment of the volcanic hazard. *Natural Hazards and Earth System Sciences*, 9(2):425–431.
- Mesimeri, M. and Karakostas, V. (2018). Repeating earthquakes in western Corinth Gulf (Greece): implications for aseismic slip near locked faults. *Geophysical Journal International*, 215(1):659–676.
- Mesimeri, M., Karakostas, V., Papadimitriou, E., Schaff, D., and Tsaklidis, G. (2016). Spatio-temporal properties and evolution of the 2013 Aigion earthquake swarm (Corinth Gulf, Greece). *Journal of Seismology*, 20(2):595–614.
- Mesimeri, M., Karakostas, V., Papadimitriou, E., and Tsaklidis, G. (2019). Characteristics of earthquake clusters: Application to western Corinth Gulf (Greece). *Tectonophysics*, 767:228160.
- Mesimeri, M., Karakostas, V., Papadimitriou, E., Tsaklidis, G., and Jacobs, K. (2018a). Relocation of recent seismicity and seismotectonic properties in the Gulf of Corinth (Greece). *Geophysical Journal International*, 212(2):1123–1142.
- Mesimeri, M., Kourouklas, C., Papadimitriou, E., Karakostas, V., and Kementzetzidou, D. (2018b). Analysis of microseismicity associated with the 2017 seismic swarm near the Aegean coast of NW Turkey. *Acta Geophysica*, 66(4):479–495.
- Michael, A. J. (2011). Random variability explains apparent global clustering of large earthquakes. *Geophysical Research Letters*, 38(21).
- Michael, A. J. (2014). How complete is the ISC-GEM global earthquake catalog? *Bulletin of the Seismological Society of America*, 104(4):1829–1837.

REFERENCES

---

- Michas, G., Kapetanidis, V., Kaviris, G., and Vallianatos, F. (2021). Earthquake Diffusion Variations in the Western Gulf of Corinth (Greece). *Pure and Applied Geophysics*, pages 1–16.
- Mignan, A. and Chouliaras, G. (2014). Fifty years of seismic network performance in Greece (1964–2013): spatiotemporal evolution of the completeness magnitude. *Seismological Research Letters*, 85(3):657–667.
- Mizrahi, L., Nandan, S., and Wiemer, S. (2021). The effect of declustering on the size distribution of mainshocks. *Seismological Research Letters*, 92(4):2333–2342.
- Molchan, G. and Dmitrieva, O. (1992). Aftershock identification: methods and new approaches. *Geophysical Journal International*, 109(3):501–516.
- Montoro-Cazorla, D., Pérez-Ocón, R., and del Carmen Segovia, M. (2009). Replacement policy in a system under shocks following a Markovian arrival process. *Reliability Engineering & System Safety*, 94(2):497–502.
- Nandan, S., Ouillon, G., Sornette, D., and Wiemer, S. (2019). Forecasting the full distribution of earthquake numbers is fair, robust, and better. *Seismological Research Letters*, 90(4):1650–1659.
- Neuts, M. F. (1978). Renewal processes of phase type. *Naval Research Logistics Quarterly*, 25(3):445–454.
- Neuts, M. F. (1979). A Versatile Markovian Point Process. *Journal of Applied Probability*, 16(4):764–779.
- O’cinneide, C. A. (1999). Phase-type distributions: open problems and a few properties. *Stochastic Models*, 15(4):731–757.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, 83(401):9–27.

REFERENCES

---

- Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2):379–402.
- Okamura, H. and Dohi, T. (2009). Faster maximum likelihood estimation algorithms for Markovian arrival processes. In *Proceedings of 6th International Conference on Quantitative Evaluation of Systems (QEST2009)*, volume 9, pages 73–82.
- Okamura, H. and Dohi, T. (2016). Fitting phase-type distributions and Markovian arrival processes: Algorithms and tools. In *Principles of performance and reliability modeling and evaluation*, pages 49–75. Springer.
- Omi, T., Ogata, Y., Hirata, Y., and Aihara, K. (2015). Intermediate-term forecasting of aftershocks from an early aftershock sequence: Bayesian and ensemble forecasting approaches. *Journal of Geophysical Research: Solid Earth*, 120(4):2561–2578.
- Ōmori, F. (1894). On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo*, 7:111–200.
- Orfanogiannaki, K. and Karlis, D. (2018). Multivariate poisson hidden markov models with a case study of modelling seismicity. *Australian New Zealand Journal of Statistics*, 60(3):301–322.
- Orfanogiannaki, K., Karlis, D., and Papadopoulos, G. (2010). Identifying seismicity levels via Poisson hidden Markov models. *Pure and Applied Geophysics*, 167(8-9):919–931.
- Orlik, P. V. and Rappaport, S. S. (1998). A model for teletraffic performance and channel holding time characterization in wireless cellular communication with general session and dwell time distributions. *IEEE Journal on Selected Areas in Communications*, 16(5):788–803.
- Page, M. T., Van Der Elst, N., Hardebeck, J., Felzer, K., and Michael, A. J. (2016). Three ingredients for improved global aftershock forecasts: Tectonic

REFERENCES

---

region, time-dependent catalog incompleteness, and intersequence variability. *Bulletin of the Seismological Society of America*, 106(5):2290–2301.

Papadimitriou, E., Karakostas, V., Mesimeri, M., Chouliaras, G., and Kourouklas, C. (2017). The Mw6.5 17 November 2015 Lefkada (Greece) earthquake: structural interpretation by means of the aftershock analysis. *Pure and Applied Geophysics*, 174(10):3869–3888.

Papadimitriou, E. E. (1993). Focal Mechanism along the Convex Side of the Hellenic Arc. *Bollettino di Geofisica Teorica ed Applicata*, 35(140):402–426.

Papadimitriou, E. E. (2002). Mode of strong earthquake recurrence in the central Ionian Islands (Greece): possible triggering due to Coulomb stress changes generated by the occurrence of previous strong shocks. *Bulletin of the Seismological Society of America*, 92(8):3293–3308.

Papadimitriou, P., Kassaras, I., Kaviris, G., Tselentis, G.-A., Voulgaris, N., Lekkas, E., Chouliaras, G., Evangelidis, C., Pavlou, K., Kapetanidis, V., et al. (2018). The 12th June 2017 Mw=6.3 Lesvos earthquake from detailed seismological observations. *Journal of Geodynamics*, 115:23–42.

Papadopoulos, G. A., Karastathis, V. K., Ganas, A., Pavlides, S., Fokaefs, A., and Orfanogiannaki, K. (2003). The Lefkada, Ionian Sea (Greece), shock (Mw6.2) of 14 August 2003 Evidence for the characteristic earthquake from seismicity and ground failures. *Earth, planets and space*, 55(11):713–718.

Papangelou, F. (1972). Integrability of expected increments of point processes and a related random change of scale. *Transactions of the American Mathematical Society*, 165:483–506.

Papazachos, B. and Comninakis, P. (1971). Geophysical and tectonic features of the Aegean arc. *Journal of Geophysical Research*, 76(35):8517–8533.

## REFERENCES

---

- Papazachos, B., Comninakis, P., Papadimitriou, E., and Scordilis, E. (1984). Properties of the February-March 1981 seismic sequence in the Alkyonides gulf in central Greece. *Annals of Geophysics*, 2:537–544.
- Papazachos, B., Papadimitriou, E., Kiratzi, A., Papazachos, C., and Louvari, E. (1998). Fault plane solutions in the Aegean Sea and the surrounding area and their tectonic implication. *Bolletino ti Geofisica Teorica ed Applicata*, 39(3):199–218.
- Papazachos, B. and Papazachou, C. (2003). *The earthquakes of Greece*. Ziti Publ., Thessaloniki.
- Parsons, T. and Geist, E. L. (2012). Were global  $M \geq 8.3$  earthquake time intervals random between 1900 and 2011? *Bulletin of the Seismological Society of America*, 102(4):1583–1592.
- Parsons, T. and Velasco, A. A. (2011). Absence of remotely triggered large earthquakes beyond the mainshock region. *Nature Geoscience*, 4(5):312–316.
- Peng, W., Chen, K. H., and Toda, S. (2019). Evaluating the association between tectonic tremors and earthquakes in Taiwan from 7 years catalogs. *Journal of Geophysical Research: Solid Earth*, 124(4):3950–3965.
- Peresan, A. and Gentili, S. (2018). Seismic clusters analysis in Northeastern Italy by the nearest-neighbor approach. *Physics of the Earth and Planetary Interiors*, 274:87–104.
- Peresan, A. and Gentili, S. (2020). Identification and characterisation of earthquake clusters: a comparative analysis for selected sequences in Italy and adjacent regions. *Bollettino di Geofisica Teorica ed Applicata*, 61(1).
- Pertsinidou, C., Tsaklidis, G., Papadimitriou, E., and Limnios, N. (2017). Application of hidden semi-Markov models for the seismic hazard assessment of the North and South Aegean Sea, Greece. *Journal of Applied Statistics*, 44(6):1064–1085.

REFERENCES

---

- Petersen, G., Niemz, P., Cesca, S., Mouslopoulou, V., and Bocchini, G. (2021). Clusty, the waveform-based network similarity clustering toolbox: concept and application to image complex faulting offshore Zakynthos (Greece). *Geophysical Journal International*, 224(3):2044–2059.
- Petersen, M. D., Mueller, C. S., Moschetti, M. P., Hoover, S. M., Rukstales, K. S., McNamara, D. E., Williams, R. A., Shumway, A. M., Powers, P. M., Earle, P. S., Llenos, A. L., Michael, A. J., Rubinstein, J. L., Norbeck, J. H., and Cochran, E. S. (2018). 2018 One-Year Seismic Hazard Forecast for the Central and Eastern United States from Induced and Natural Earthquakes. *Seismological Research Letters*, 89(3):1049–1061.
- Pievatolo, A. and Rotondi, R. (2008). Statistical identification of seismicity phases. *Geophys. J. Int.*, 173(3):942–957.
- Pollitz, F. F., Stein, R. S., Sevilgen, V., and Bürgmann, R. (2012). The 11 April 2012 east Indian Ocean earthquake triggered large aftershocks worldwide. *Nature*, 490(7419):250–253.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286.
- Reasenber, P. (1985). Second-order moment of central California seismicity, 1969–1982. *Journal of Geophysical Research: Solid Earth*, 90(B7):5479–5495.
- Reasenber, P. and Jones, L. (1994). Earthquake aftershocks: update. *Science*, 265(5176):1251–1252.
- Rhoades, D. A., Christophersen, A., Gerstenberger, M. C., Liukis, M., Silva, F., Marzocchi, W., Werner, M. J., and Jordan, T. H. (2018). Highlights from the first ten years of the New Zealand earthquake forecast testing center. *Seismological Research Letters*, 89(4):1229–1237.

REFERENCES

---

- Rhoades, D. A., Schorlemmer, D., Gerstenberger, M. C., Christophersen, A., Zechar, J. D., and Imoto, M. (2011). Efficient testing of earthquake forecasting models. *Acta Geophysica*, 59(4):728–747.
- Rigo, A., Lyon-Caen, H., Armijo, R., Deschamps, A., Hatzfeld, D., Makropoulos, K., Papadimitriou, P., and Kassaras, I. (1996). A microseismic study in the western part of the Gulf of Corinth (Greece): Implications for large-scale normal faulting mechanisms. *Geophysical Journal International*, 126(3):663–688.
- Rogerson, P. A. (2018). Statistical evidence for long-range space-time relationships between large earthquakes. *Journal of Seismology*, 22(6):1423–1435.
- Ross, Z. E., Trugman, D. T., Hauksson, E., and Shearer, P. M. (2019). Searching for hidden earthquakes in Southern California. *Science*, 364(6442):767–771.
- Rydén, T. (1996). An EM algorithm for estimation in Markov-modulated Poisson processes. *Computational Statistics & Data Analysis*, 21(4):431–447.
- Saltogianni, V., Gianniou, M., Taymaz, T., Yolsal-Çevikbilen, S., and Stiros, S. (2015). Fault slip source models for the 2014 Mw6.9 Samothraki-Gökçeada earthquake (North Aegean trough) combining geodetic and seismological observations. *Journal of Geophysical Research: Solid Earth*, 120(12):8610–8622.
- Sammis, C. G. and Smith, S. W. (2013). Triggered tremor, phase-locking, and the global clustering of great earthquakes. *Tectonophysics*, 589:167–171.
- Savran, W. H., Werner, M. J., Marzocchi, W., Rhoades, D. A., Jackson, D. D., Milner, K., Field, E., and Michael, A. (2020). Pseudoprospective evaluation of UCERF3-ETAS forecasts during the 2019 Ridgecrest sequence. *Bulletin of the Seismological Society of America*, 110(4):1799–1817.

REFERENCES

---

- Schneider, L. F., Krajina, A., and Krivobokova, T. (2021). Threshold selection in univariate extreme value analysis. *Extremes*, 24(4):881–913.
- Scholz, C. H. (2010). Large earthquake triggering, clustering, and the synchronization of faults. *Bulletin of the Seismological Society of America*, 100(3):901–909.
- Scholz, C. H. (2019). *The mechanics of earthquakes and faulting*. Cambridge university press.
- Schorlemmer, D. and Gerstenberger, M. (2007). RELM testing center. *Seismological Research Letters*, 78(1):30–36.
- Schorlemmer, D., Gerstenberger, M., Wiemer, S., Jackson, D., and Rhoades, D. (2007). Earthquake likelihood model testing. *Seismological Research Letters*, 78(1):17–29.
- Schorlemmer, D., Werner, M. J., Marzocchi, W., Jordan, T. H., Ogata, Y., Jackson, D. D., Mak, S., Rhoades, D. A., Gerstenberger, M. C., Hirata, N., et al. (2018). The collaboratory for the study of earthquake predictability: Achievements and priorities. *Seismological Research Letters*, 89(4):1305–1313.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.
- Scordilis, E., Karakaisis, G., Karacostas, B., Panagiotopoulos, D., Comninakis, P., and Papazachos, B. (1985). Evidence for transform faulting in the Ionian Sea: the Cephalonia island earthquake sequence of 1983. *Pure and Applied Geophysics*, 123(3):388–397.
- Seif, S., Mignan, A., Zechar, J. D., Werner, M. J., and Wiemer, S. (2017). Estimating ETAS: The effects of truncation, missing data, and model assumptions. *Journal of Geophysical Research: Solid Earth*, 122(1):449–469.

REFERENCES

---

- Shearer, P. M. and Stark, P. B. (2012). Global risk of big earthquakes has not recently increased. *Proceedings of the National Academy of Sciences*, 109(3):717–721.
- Shebalin, P. N., Narteau, C., and Baranov, S. V. (2020). Earthquake productivity law. *Geophysical Journal International*, 222(2):1264–1269.
- Sheikhhosseini, Z., Mirzaei, N., Heidari, R., and Monkaresi, H. (2021). Delineation of potential seismic sources using weighted K-means cluster analysis and particle swarm optimization (PSO). *Acta Geophysica*, 69(6):2161–2172.
- Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, pages 367–377.
- Sobradelo, R., Martí, J., Mendoza-Rosas, A., and Gómez, G. (2011). Volcanic hazard assessment for the Canary Islands (Spain) using extreme value theory. *Natural Hazards and Earth System Sciences*, 11(10):2741–2753.
- Sokos, E., Kiratzi, A., Gallovič, F., Zahradník, J., Serpetsidaki, A., Plicka, V., Janský, J., Kostelecký, J., and Tselentis, G.-A. (2015). Rupture process of the 2014 Cephalonia, Greece, earthquake doublet (Mw6) as inferred from regional and local seismic data. *Tectonophysics*, 656:131–141.
- Sokos, E., Zahradník, J., Kiratzi, A., Janský, J., Gallovič, F., Novotny, O., Kostelecký, J., Serpetsidaki, A., and Tselentis, G.-A. (2012). The January 2010 Efpalio earthquake sequence in the western Corinth gulf (Greece). *Tectonophysics*, 530:299–309.
- Stallone, A. and Marzocchi, W. (2019). Features of Seismic Sequences Are Similar in Different Crustal Tectonic Regions. *Bulletin of the Seismological Society of America*, 109(5):1594–1604.
- Stein, R. S. (1999). The role of stress transfer in earthquake occurrence. *Nature*, 402(6762):605–609.

REFERENCES

---

- Stirling, M. W., Verry, G. H. M., and Berryman, K. R. (2002). A new seismic hazard model for New Zealand. *Bulletin of the Seismological Society of America*, 92(5):1878–1903.
- Storchak, D. A., Di Giacomo, D., Bondár, I., Engdahl, E. R., Harris, J., Lee, W. H., Villaseñor, A., and Bormann, P. (2013). Public release of the ISC–GEM global instrumental earthquake catalogue (1900–2009). *Seismological Research Letters*, 84(5):810–815.
- Storchak, D. A., Di Giacomo, D., Engdahl, E., Harris, J., Bondár, I., Lee, W. H., Bormann, P., and Villaseñor, A. (2015). The ISC-GEM global instrumental earthquake catalogue (1900–2009): introduction. *Physics of the Earth and Planetary Interiors*, 239:48–63.
- Tahir, M., Grasso, J.-R., and Amorèse, D. (2012). The largest aftershock: How strong, how far away, how delayed? *Geophysical Research Letters*, 39(4).
- Taroni, M. and Akinci, A. (2021). Good practices in PSHA: declustering, b-value estimation, foreshocks and aftershocks inclusion; a case study in Italy. *Geophysical Journal International*, 224(2):1174–1187.
- Taymaz, T., Jackson, J., and McKenzie, D. (1991). Active tectonics of the north and central Aegean Sea. *Geophysical Journal International*, 106(2):433–490.
- Telek, M. and Horváth, G. (2007). A minimal representation of markov arrival processes and a moments matching method. *Performance Evaluation*, 64(9–12):1153–1168.
- Teng, G. and Baker, J. W. (2019). Seismicity Declustering and Hazard Analysis of the Oklahoma–Kansas Region. *Bulletin of the Seismological Society of America*, 109(6):2356–2366.

REFERENCES

---

- Touati, S., Naylor, M., and Main, I. (2016). Detection of change points in underlying earthquake rates, with application to global mega-earthquakes. *Geophysical Journal International*, 204(2):753–767.
- Towe, R., Tawn, J., Eastoe, E., and Lamb, R. (2020). Modelling the clustering of extreme events for short-term risk assessment. *Journal of Agricultural, Biological and Environmental Statistics*, 25(1):32–53.
- University of Athens (2008). Hellenic Seismological Network, University of Athens, Seismological Laboratory. International Federation of Digital Seismograph Networks. dataset available at <http://dx.doi.org/doi:10.7914/SN/HA>.
- Utsu, T. (1971). Aftershocks and earthquake statistics (2): further investigation of aftershocks and other earthquake sequences based on a new classification of earthquake sequences. *Journal of the Faculty of Science, Hokkaido University. Series 7, Geophysics*, 3(4):197–266.
- Utsu, T., Ogata, Y., et al. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth*, 43(1):1–33.
- van Stiphout, T., Zhuang, J., and Marsan, D. (2012). Seismicity declustering. *Community Online Resource for Statistical Seismicity Analysis*, 10(1).
- Varini, E., Peresan, A., and Zhuang, J. (2020). Topological comparison between the stochastic and the nearest-neighbor earthquake declustering methods through network analysis. *Journal of Geophysical Research: Solid Earth*, 125(8):e2020JB019718.
- Vere-Jones, D. (2010). Foundations of statistical seismology. *Pure and applied geophysics*, 167(6):645–653.
- Vijay, R. K. and Nanda, S. J. (2021). Seismicity analysis using space-time density peak clustering method. *Pattern Analysis and Applications*, 24(1):181–201.

REFERENCES

---

- Votsi, I., Limnios, N., Tsaklidis, G., and Papadimitriou, E. (2013). Hidden markov models revealing the stress field underlying the earthquake generation. *Physica A*, 392(13):2868–2885.
- Votsi, I., Limnios, N., Tsaklidis, G., and Papadimitriou, E. (2014). Hidden semi-Markov modeling for the estimation of earthquake occurrence rates. *Communications in Statistics-Theory and Methods*, 43(7):1484–1502.
- Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods*, 17(2):228.
- Wang, Q., Jackson, D. D., and Zhuang, J. (2010). Missing links in earthquake clustering models. *Geophysical Research Letters*, 37(21).
- Wang, T., Bebbington, M., and Harte, D. (2012). Markov-modulated Hawkes process with stepwise decay. *Annals of the Institute of Statistical Mathematics*, 64(3):521–544.
- Wang, T., Schofield, M., Bebbington, M., and Kiyosugi, K. (2020). Bayesian modelling of marked point processes with incomplete records: volcanic eruptions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 69(1):109–130.
- Wang, T., Zhuang, J., Buckby, J., Obara, K., and Tsuruoka, H. (2018). Identifying the Recurrence Patterns of Nonvolcanic Tremors Using a 2-D Hidden Markov Model With Extra Zeros. *Journal of Geophysical Research: Solid Earth*, 123(8):6802–6825.
- Wang, T., Zhuang, J., Obara, K., and Tsuruoka, H. (2017). Hidden Markov modelling of sparse time series from non-volcanic tremor observations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 66(4):691–715.

REFERENCES

---

- Wells, D. L. and Coppersmith, K. J. (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.
- Wiemer, S. (2001). A software package to analyze seismicity: ZMAP. *Seismological Research Letters*, 72(3):373–382.
- Wiemer, S. and Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bulletin of the Seismological Society of America*, 90(4):859–869.
- Wu, Z. (2010). A hidden Markov model for earthquake declustering. *Journal of Geophysical Research: Solid Earth*, 115(B3).
- Wu, Z. (2011). Quasi-hidden Markov model and its applications in cluster analysis of earthquake catalogs. *Journal of Geophysical Research: Solid Earth*, 116(B12).
- Yamashita, T. (1999). Pore creation due to fault slip in a fluid-permeated fault zone and its effect on seismicity: generation mechanism of earthquake swarm. In *Seismicity patterns, their statistical significance and physical meaning*, pages 625–647. Springer.
- Yip, C. F., Ng, W. L., and Yau, C. Y. (2018). A hidden Markov model for earthquake prediction. *Stochastic Environmental Research and Risk Assessment*, 32(5):1415–1434.
- Zaliapin, I. and Ben-Zion, Y. (2013a). Earthquake clusters in southern California I: Identification and stability. *Journal of Geophysical Research: Solid Earth*, 118(6):2847–2864.
- Zaliapin, I. and Ben-Zion, Y. (2013b). Earthquake clusters in southern California II: Classification and relation to physical properties of the crust. *Journal of Geophysical Research: Solid Earth*, 118(6):2865–2877.

REFERENCES

---

Zaliapin, I. and Ben-Zion, Y. (2016). A global classification and characterization of earthquake clusters. *Geophysical Journal International*, 207(1):608–634.

Zaliapin, I. and Ben-Zion, Y. (2020). Earthquake declustering using the nearest-neighbor approach in space-time-magnitude domain. *Journal of Geophysical Research: Solid Earth*, 125(4):e2018JB017120.

Zaliapin, I. and Ben-Zion, Y. (2022). Perspectives on clustering and declustering of earthquakes. *Seismological Society of America*, 93(1):386–401.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical review letters*, 101(1):018501.

Zaliapin, I. and Kreemer, C. (2017). Systematic fluctuations in the global seismic moment release. *Geophysical Research Letters*, 44(10):4820–4828.

Zechar, J. D. and Jordan, T. H. (2008). Testing alarm-based earthquake predictions. *Geophysical Journal International*, 172(2):715–724.

Zhang, Q. and Shearer, P. M. (2016). A new method to identify earthquake swarms applied to seismicity near the San Jacinto Fault, California. *Geophysical Journal International*, 205(2):995–1005.

Zhuang, J., Chang, C.-P., Ogata, Y., and Chen, Y.-I. (2005). A study on the background and clustering seismicity in the Taiwan region by using point process models. *Journal of Geophysical Research: Solid Earth*, 110(B5).

Zhuang, J., Harte, D., Werner, M., Hainzl, S., and Zhou, S. (2012). Basic models of seismicity: Temporal models. Community Online Resource for Statistical Seismicity Analysis [Online]. Available: [10.5078/corssa-79905851](https://doi.org/10.5078/corssa-79905851).

Zhuang, J., Murru, M., Falcone, G., and Guo, Y. (2019). An extensive study

## REFERENCES

---

of clustering features of seismicity in Italy from 2005 to 2016. *Geophysical Journal International*, 216(1):302–318.

Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, 97(458):369–380.

Zhuang, J., Ogata, Y., and Vere-Jones, D. (2004). Analyzing earthquake clustering features by using stochastic reconstruction. *Journal of Geophysical Research: Solid Earth*, 109(B5).