

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΤΟΥ ΦΑΙΝΟΜΕΝΟΥ
ΤΣΟΥΝΑΜΙ ΛΟΓΩ ΣΕΙΣΜΩΝ**

Πέτρος Α. Τσακμάκης

Διπλωματική Εργασία
που υποβλήθηκε στο Τμήμα Στατιστικής και
Ασφαλιστικής Επιστήμης του Πανεπιστημίου
Πειραιώς ως μέρος των απαιτήσεων για την
απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην Εφαρμοσμένη Στατιστική.

Πειραιάς
Μάιος 2015

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΤΟΥ ΦΑΙΝΟΜΕΝΟΥ
ΤΣΟΥΝΑΜΙ ΛΟΓΩ ΣΕΙΣΜΩΝ**

Πέτρος Α. Τσακμάκης

Διπλωματική Εργασία
που υποβλήθηκε στο Τμήμα Στατιστικής και
Ασφαλιστικής Επιστήμης του Πανεπιστημίου
Πειραιώς ως μέρος των απαιτήσεων για την
απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην Εφαρμοσμένη Στατιστική.

Πειραιάς
Μάιος 2015

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική.

Τα μέλη της Επιτροπής ήταν:

- Επίκουρος Καθηγητής Τζαβελάς Γεώργιος (Επιβλέπων)
- Επίκουρος Καθηγητής Ψαρράκος Γεώργιος
- Αναπληρωτής Καθηγητής Πολίτης Κωνσταντίνος

Η έγκριση της Διπλωματική Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνωμών του συγγραφέα.

UNIVERSITY OF PIRAEUS



**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**STATISTICAL ANALYSIS OF TSUNAMI
PHENOMENON DUE TO EARTHQUAKE**

By

Petros A. Tsakmakis

MSc Dissertation

submitted to the Department of Statistics and
Insurance Science of the University of Piraeus in
partial fulfilment of the requirements for the degree
of Master of Science in Applied Statistics.

Piraeus, Greece

May 2015

“Statistics is the grammar of science”

Karl Pearson

Ευχαριστίες

Θέλω να ευχαριστήσω τους γονείς μου για την υποστήριξη τους σε όλη την προσπάθεια των σπουδών μου. Τον κ. Καλογήρου για την βοήθεια του στο πρόγραμμα GIS καθώς επίσης τον κ. Μπερσίμη Σωτήρη για την όλη βοήθεια και υπομονή στις απορίες μου. Επίσης την Ευτυχία Κουρινάκη για την πολύτιμη βοήθεια της στην κατασκευή και αντιπαραβολή των δεδομένων που χρειάστηκαν για την εκπόνηση αυτής της διπλωματικής. Τέλος θα ευχαριστήσω όλους τους συμφοιτητές καθώς και καθηγητές μου που είτε με τη συμπαράσταση είτε με την αμφισβήτησή τους συνέβαλαν στο να γίνω καλύτερος.

Περίληψη

Στην παρούσα διπλωματική γίνεται η στατιστική ανάλυση του φαινομένου του τσουνάμι. Διερευνήθηκαν και αξιολογήθηκαν μοντέλα πρόβλεψης του ύψους κύματος του τσουνάμι με βάση τα χαρακτηριστικά των σεισμών καθώς επίσης και η πιθανότητα εμφάνισης τσουνάμι λόγω σεισμικής δραστηριότητας. Επίσης διερευνήθηκε το κατάλληλο μέγεθος σεισμού που είναι ικανό για τη δημιουργία τσουνάμι. Τέλος διερευνήθηκαν περιοχές ή ακόμα και χώρες που διατρέχουν μεγαλύτερο κίνδυνο εμφάνισης τσουνάμι από τον υπόλοιπο κόσμο και έγινε μια προσπάθεια μελέτης της περιοδικότητας του φαινομένου όσον αφορά την Ελλάδα.

Abstract

In the present MSc Dissertation we performed the statistical analysis of the tsunami phenomenon. Forecasting models for the height of tsunami wave were investigated and evaluated from basic earthquake characteristics as well as for the probability of occurrence of tsunami event due to seismic activity. Moreover we investigated the earthquake magnitude capable to cause tsunami. Finally, we analyzed the risk of a tsunami appearance in several countries or even larger geographical areas and an effort was made to study the periodicity of the phenomenon as far as it concerns Greece.

Περιεχόμενα

Περιεχόμενα	XV
Κατάλογος Πινάκων	XVII
Κατάλογος Εικόνων.....	XVIII

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή - Ορισμοί και έννοιες σεισμολογικών φαινομένων

1.1 Εισαγωγή	1
1.2 Σεισμοί.....	1
1.2.1 Γένεση σεισμών.....	1
1.2.2 Μέτρηση Σεισμών.....	2
1.2.2.1 Μέγεθος.....	2
1.2.2.2 Ένταση.....	4
1.3 Τσουνάμι.....	5
1.3.1 Ταχύτητα.....	5
1.3.2 Ύψος κύματος.....	5
1.3.3 Μέγεθος.....	6
1.3.4 Ένταση.....	6

ΚΕΦΑΛΑΙΟ 2: Στατιστικές Μέθοδοι Ανάλυσης Φαινομένων Τσουνάμι

2.1 Ανάλυση Παλινδρόμησης.....	8
2.2 Ανάλυση Bootstrap.....	10
2.3 Λογιστική Παλινδρόμηση.....	10
2.4 Διατακτική Παλινδρόμηση.....	11
2.5 Discriminant Analysis.....	12
2.6 Χωρική Ανάλυση.....	13

ΚΕΦΑΛΑΙΟ 3: Εφαρμογή Στατιστικών Αναλύσεων σε Παγκόσμια Δεδομένα

3.1	Περιγραφή μεταβλητών.....	16
3.2	Περιγραφική Στατιστική Ανάλυση.....	18
3.3	Γεωχωρική Ανάλυση.....	22
3.4	Ανάλυση Παλινδρόμησης.....	24
3.5	Λογιστική Παλινδρόμηση.....	26
3.6	Διατακτική Παλινδρόμηση.....	30
3.7	Discriminant Analysis.....	34
3.8	Χωρική Ανάλυση.....	47
3.9	Περιοδικότητα.....	60

ΚΕΦΑΛΑΙΟ 4: Εμπόδια εκπόνησης και συμπεράσματα

4.1	Εμπόδια εκπόνησης.....	63
4.2	Συμπεράσματα-Προτάσεις.....	64

Παράρτημα Α	65
--------------------------	-----------

Βιβλιογραφία	72
---------------------------	-----------

Κατάλογος πινάκων

1.1	Τροποποιημένη Mercalli	4
1.2	Κλίμακα έντασης τσουνάμι	6
3.1	Συχνότητες σεισμών ανά χώρα	18
3.2	Συχνότητες σεισμών ανά ευρύτερη γεωγραφική περιοχή	19
3.3	Περιγραφικά στοιχεία σεισμών ανά χώρα	20
3.4	Παλινδρόμηση ανά γεωγραφική περιοχή	24
3.5	Παλινδρόμηση ανά χώρα	26
3.6	Λογιστική παλινδρόμηση ανά γεωγραφική περιοχή	27
3.7	Λογιστική παλινδρόμηση ανά χώρα	29
3.8	Πίνακας συχνοτήτων επιπέδων έντασης Τσουνάμι.....	31
3.9	Omnibus test διατακτικής παλινδρόμησης.....	31
3.10	Μέτρα καλής προσαρμογής.....	31
3.11	Δείκτες ψευδο- R^2	32
3.12	Πίνακας συντελεστών διατακτικής παλινδρόμησης.....	32
3.13	Έλεγχος παράλληλων γραμμών.....	33
3.14	AUC ανά γεωγραφική περιοχή-Μέγεθος σεισμού	34
3.15	Ποσοστά ορθής ταξινόμησης ανά γεωγραφική περιοχή-Μέγεθος σεισμού	35
3.16	AUC ανά χώρα-Μέγεθος σεισμού	36
3.17	Ποσοστά ορθής ταξινόμησης ανά χώρα-Μέγεθος σεισμού	37
3.18	AUC ανά γεωγραφική περιοχή-Ένταση σεισμού	38
3.19	Ποσοστά ορθής ταξινόμησης ανά γεωγραφική περιοχή-Ένταση σεισμού	38
3.20	AUC ανά χώρα-Ένταση σεισμού	39
3.21	Ποσοστά ορθής ταξινόμησης ανά χώρα-Ένταση σεισμού	39
3.22	AUC επιφανειακών σεισμών ανά γεωγραφική περιοχή-Μέγεθος σεισμού.....	40
3.23	Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά γεωγραφική περιοχή-Μέγεθος σεισμού	40
3.24	AUC επιφανειακών σεισμών ανά χώρα-Μέγεθος σεισμού	41

3.25	Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά χώρα-Μέγεθος σεισμού	42
3.26	AUC επιφανειακών σεισμών ανά χώρα-Ένταση σεισμού	43
3.27	Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά χώρα-Ένταση σεισμού	43
3.28	AUC επιφανειακών σεισμών ανά γεωγραφική περιοχή-Ένταση σεισμού	43
3.29	Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά γεωγραφική περιοχή-Ένταση σεισμού	44
3.30	Πίνακας συνάφειας καταγραφών τσουνάμι με την αξιολόγηση τους	44
3.31	Συστάδες ανά χώρα	53
3.32	Συστάδες ανά γεωγραφική περιοχή	59

Κατάλογος εικόνων

2.1	Διάγραμμα διασποράς παράδειγμα	9
3.1	Χάρτης σεισμών	22
3.2	Χάρτης τσουνάμι	23
3.3	Γράφημα αξιοπιστίας συνόλου σεισμών	45
3.4	Γράφημα αξιοπιστίας επιφανειακών σεισμών	46
3.5	Γράφημα περιοδικότητας συνόλου τσουνάμι	60
3.6	Γράφημα περιοδικότητας επιβεβαιωμένων τσουνάμι	61
3.7	Γράφημα περιοδικότητας σεισμικής δραστηριότητας	62

Κεφάλαιο 1.

Εισαγωγή – Ορισμοί και έννοιες σεισμολογικών φαινομένων

1.1 Εισαγωγή

Οι συνέπειες ενός σεισμού καθώς και των φαινομένων τσουνάμι είναι στην πλειοψηφία τους καταστροφικές και θανατηφόρες. Η μελέτη και πρόβλεψη τέτοιων φαινομένων κρίνονται υψηλής σημαντικότητας. Η πρόβλεψη σεισμών έχει προσεγγιστεί από αρκετούς επιστήμονες είτε ντετερμινιστικά είτε στοχαστικά με μοντέλα της θεωρίας ακραίων τιμών. Παρόλα αυτά η πρόβλεψη των φαινομένων τσουνάμι, στη διεθνή βιβλιογραφία, δεν έχει προσεγγιστεί από κάποιο ικανοποιητικό στοχαστικό μοντέλο. Σκοπός της παρούσας μελέτης είναι η διερεύνηση ενός τέτοιου μοντέλου καθώς και όλες οι πιθανές μεθοδολογίες για την στατιστική ανάλυση των φαινομένων τσουνάμι. Πριν την αναλυτική παρουσίαση τέτοιων μεθόδων είναι αναπόφευκτη η εισαγωγή σε ορισμούς και έννοιες που χρησιμοποιούνται στην περιγραφή σεισμών και τσουνάμι.

1.2 Σεισμοί

1.2.1 Γένεση σεισμών

Ο κύριος λόγος γένεσης σεισμών είναι η κίνηση των λιθοσφαιρικών πλακών κάτω από το φλοιό της Γης. Οι κινήσεις των λιθοσφαιρικών πλακών είναι είτε απόκλισης είτε σύγκλισης είτε παράλληλες. Κατά την κίνηση αυτή αναπτύσσονται δυνάμεις τριβής στα πετρώματα της λιθόσφαιρας και όταν αυτές ξεπεράσουν το σημείο ισορροπίας, τα πετρώματα θραύονται. Η θραύση των πετρωμάτων έχει ως αποτέλεσμα τη δημιουργία ρηγμάτων. Οι πλευρές των ρηγμάτων, προφανώς, δεν είναι ομαλές και οι δυνάμεις, ακόμα και σε κατάσταση ισορροπίας, που ασκούνται στις πλευρές των ρηγμάτων προκαλούν ταλάντωση των σημείων εφαρμογής των δυνάμεων αυτών. Η ταλάντωση των σημείων είναι αλυσιδωτή και λειτουργεί με τη μορφή ντόμινο σε όλα τα γειτονικά σημεία των πλευρών του ρήγματος. Οι ταλαντώσεις μεταδίδονται έτσι μέσα στη Γη (*σεισμικά κύματα*) και όταν φτάνουν στην επιφάνεια του φλοιού της προκαλούν τα φαινόμενα των σεισμών.

Οι σεισμοί χωρίζονται σε τρεις κατηγορίες με βάση το βάθος γένεσης της ταλάντωσης (*εστιακό βάθος*). Σεισμοί με εστιακό βάθος μικρότερο των 60 χιλιομέτρων

ονομάζονται *επιφανειακοί*, σεισμοί με εστιακό βάθος μεταξύ 60 και 300 χιλιομέτρων ονομάζονται *ενδιάμεσων βάθους* ενώ σεισμοί με εστιακό βάθος άνω των 300 χιλιομέτρων ονομάζονται *βάθους*.

1.2.2 Μέτρηση σεισμών

Η μέτρηση της ενέργειας ενός σεισμού γίνεται με το μέγεθος (M) ενός σεισμού. Επίσης η ένταση (I) χρησιμοποιείται για την εκτίμηση των μακροσεισμικών αποτελεσμάτων (βλάβες δρόμων και κτιρίων κλπ) ενός σεισμού. Η ένταση διαφέρει από τόπο σε τόπο ανάλογα με την απόσταση από το επίκεντρο καθώς και από το υλικό του φλοιού της Γης. Με άλλα λόγια σεισμοί ίδιου μεγέθους προκαλούν διαφορετικές ζημιές στο περιβάλλον.

1.2.2.1 Μέγεθος (Magnitude)

Το μέγεθος M είναι το μέτρο της ενέργειας (δυναμικής) που απελευθερώνεται με τη μορφή ταλάντωσης του φλοιού της Γης. Ο υπολογισμός του γίνεται με την μέτρηση ενός διανύσματος σεισμικών παραμέτρων (πλάτος σεισμικού κύματος, διάρκεια σεισμού, περίοδος σεισμικού κύματος) επί των σειсмоγραφημάτων. Οι κυματικές μορφές που χρησιμοποιούνται ποικίλουν με αποτέλεσμα να υπάρχουν διάφορες κλίμακες μεγεθών. Έτσι, για διαφορετικές αποστάσεις από το επίκεντρο του σεισμού ο υπολογισμός του μεγέθους βασίζεται σε διαφορετικά σεισμικά κύματα. Οι κλίμακες αυτές από τον ορισμό τους δεν έχουν θεωρητικά ανώτατο ή κατώτατο όριο, όμως όρια θέτονται από τη φυσική αντοχή των πετρωμάτων του φλοιού.

Τοπικό Μέγεθος M_L (Local Magnitude)

Ονομάζεται ο δεκαδικός λογάριθμος του μέγιστου πλάτους καταγραφής σεισμού σε μικρόμετρα, από πρότυπο σεισμόμετρο στρέψης βραχείας περιόδου, με απόσταση 100 χιλιομέτρων από το επίκεντρο του σεισμού και παρουσιάστηκε από τον Charle Richter το 1935. Η επικρατούσα ονομασία τέτοιων σεισμομέτρων είναι Wood-Anderson ενώ η κλίμακα αυτή είναι σημαντική για τους μηχανικούς καθώς μεγεθύνει σεισμικά κύματα που οι περιόδοί τους (0,5-1,5 δευτερόλεπτα) συμπίπτουν με την περίοδο πολλών κατασκευών.

Επιφανειακό Μέγεθος M_S (Surface Magnitude)

Χρησιμοποιείται για τον υπολογισμό του μεγέθους επιφανειακών σεισμών και βασίζεται στο πλάτος καταγραφής της ταλάντωσης του φλοιού, με περίοδο 18-22 δευτερολέπτων. Καταγράφεται από οποιοδήποτε σειсмоγράφο και σε οποιαδήποτε απόσταση από το επίκεντρο. Το M_S είναι μεγαλύτερο από το M_L , χρησιμοποιείται στην Ελλάδα και προτάθηκε από τον κ. Βασίλειο Παπαζάχο.

Χωρικό Μέγεθος M_B (Body Wave Magnitude)

Είναι μια επέκταση της κλίμακας Richter. Χρησιμοποιείται για τον υπολογισμό του μεγέθους σεισμού οποιασδήποτε απόστασης από το επίκεντρο και οποιουδήποτε εστιακού βάθους. Ο υπολογισμός βασίζεται στη μέτρηση των πλατών των επιμήκη κυμάτων (Primaes κύματα). Τα επιμήκη κύματα είναι κύματα που σχετίζονται με την μεταβολή του όγκου ή της πυκνότητας του μέσου διάδοσης. Είναι αξιόπιστο μέγεθος σεισμών με μεγαλύτερα εστιακά βάθη και για μεγάλες αποστάσεις από το επίκεντρο.

Οι παραπάνω κλίμακες έχουν βελτιωθεί με τη πάροδο των χρόνων λόγω των μειονεκτημάτων που παρουσιάζουν, το σημαντικότερο εκ των οποίων είναι ότι έχουν ανώτατο όριο τα 8 Richter. Η βελτίωση αυτή έγινε με την εισαγωγή εννοιών όπως η σεισμική ροπή από τον Ιάπωνα Hiroo Kanamori το 1979. Από τότε και μετά χρησιμοποιούνται κλίμακες μεγέθους πιο πολύπλοκες στον υπολογισμό αλλά ταυτόχρονα πιο ακριβείς όπως η κλίμακα σεισμικής ροπής M_w , η οποία συμφωνεί με την M_L για σεισμούς μικρότερους των 6 Richter και με την M_S για σεισμούς μεταξύ 6 και 8 Richter, και την M_E (Choy and Boatwright 1995), η οποία εκφράζει το δυναμικό καταστροφικότητας ενός σεισμού και χρησιμοποιείται για την ποσοτικοποίηση εκλυόμενης σεισμικής ενέργειας μεγάλων συμβάντων.

Η παρουσίαση των σχέσεων των παραπάνω κλιμάκων καθώς και του τύπου της σεισμικής ροπής μπορούν να βρεθούν στις σημειώσεις του Εργαστηρίου Αντισεισμικής Τεχνολογίας των κυρίων Κ. Σπυράκου και Ε. Τουτουδάκη και στις σημειώσεις του μαθήματος “Εισαγωγή στη Σεισμολογία” του Τομέα Γεωφυσικής του Α.Π.Θ.

1.2.2.2 Ένταση (Intensity)

Η ένταση του σεισμού είναι το μέγεθος το οποίο μετράει μακροσεισμικά τη σφοδρότητα ενός σεισμού. Σε αντίθεση με το μέγεθος του σεισμού η ένταση διαφέρει από τόπο σε τόπο καθώς και ανάλογα με την απόσταση από το επίκεντρο του σεισμού. Η μορφολογία του εδάφους καθώς και το υλικό των πετρωμάτων είναι οι σημαντικότεροι παράγοντες που επηρεάζουν την ένταση και τούτο διότι διαφέρει το πλάτος της ταλάντωσης στο μέσο μετάδοσης του κύματος.

Η πιο δημοφιλής κλίμακα μέτρησης της έντασης είναι η τροποποιημένη Mercalli (Modified Mercalli).

Πίνακας 1.1 Τροποποιημένη Mercalli

Βαθμοί Τροποποιημένης Κλίμακας Mercalli	
I (0-3 R)	Δεν γίνεται αισθητός.
II (3-3.9 R)	Αισθητός από μερικούς ανθρώπους που βρίσκονται σε ανάπαυση στους ψηλότερους ορόφους κτιρίων. Απαλά αναρτημένα αντικείμενα μπορεί να ταλαντεύονται.
III (3-3.9 R)	Αισθητός μέσα στα σπίτια. Μπορεί να μην αναγνωριστεί ως σεισμός. Δονήσεις σαν να περνάει ελαφρύ φορτηγό.
IV (4-4.9 R)	Τίθενται σε κίνηση κρεμασμένα αντικείμενα. Τζάμια τρίζουν. Σταματημένα αυτοκίνητα κλυδωνίζονται. Δονήσεις σαν να περνά βαρύ φορτηγό. Κρότος παραθύρων, χτύπος στις πόρτες.
V (4-4.9 R)	Αισθητός από σχεδόν όλους, πολλοί μπορεί να ξυπνήσουν. Μερικά πιάτα, παράθυρα, κλπ. σπάνε, σε μερικές περιπτώσεις ραγίζουν οι σοβάδες, ασταθή αντικείμενα ανατρέπονται. Μερικές φορές παρατηρείται διαταραχή των δέντρων, των στύλων, καθώς και άλλων ψηλών αντικειμένων.
VI (5-5.9 R)	Αισθητός από όλους, πολλοί φοβούνται και τρέχουν σε εξωτερικούς χώρους. Κάποια βαριά έπιπλα μετακινούνται, μερικές περιπτώσεις πεσμένων σοβάδων ή κατεστραμμένων καμινάδων. Ελαφριές ζημιές.
VII (5-5.9 R)	Μεγάλες καμπάνες ηχούν. Πτώση πολυάριθμων κεραμιδιών, καπνοδόχων. Σοβάδες και τοιχοποιία ραγίζουν στις συνηθισμένες κατασκευές. Στις κακές κατασκευές πέφτουν σοβάδες, αποκολλώνται τούβλα και πέτρες. Γίνεται αισθητός από οδηγούς αυτοκινήτων. Κυματισμός σε λίμνες, θόλωμα νερού από λάσπη.
VIII (6-6.9 R)	Επηρεάζεται η οδήγηση των αυτοκινήτων. Αρκετές ζημιές και μερική κατάρρευση στις συνηθισμένες κατασκευές. Λίγες βλάβες στην τοιχοποιία των καλών κατασκευών και μερικές στις κακές κατασκευές. Κλαδιά σπάνε από τα δέντρα. Αλλαγές στη ροή και στη θερμοκρασία του νερού σε πηγές και πηγάδια.
IX (6-6.9 R)	Γενική καταστροφή στις κακές κατασκευές. Σοβαρές βλάβες στην τοιχοποιία των καλών κατασκευών. Υπόγειοι αγωγοί σπάζουν. Υπόγειες σωληνώσεις σπασμένες.
X (≥7 R)	Καταστροφή μερικών καλών κατασκευασμένων ξύλινων κτιρίων και γεφυρών. Οι περισσότερες κατασκευές τοιχοποιίας και τα προκατασκευασμένα κτίσματα καταστρέφονται από τα θεμέλια. Σοβαρές ζημιές σε φράγματα και αναχώματα. Μεγάλες κατολισθήσεις. Οι σιδηροτροχιές κάμπτονται.
XI (≥7 R)	Μεγάλες ρωγμές στο έδαφος. Οι σιδηροτροχιές κάμπτονται έντονα. Υπόγειοι αγωγοί καταστρέφονται εντελώς.

XII (≥7 R)	Ολική καταστροφή. Παρατηρούνται κύματα στην επιφάνεια του εδάφους. Παραμόρφωση του ορίζοντα. Αντικείμενα εκτινάσσονται στον αέρα προς τα πάνω.
-----------------------------	--

1.3 Τσουνάμι

Το τσουνάμι είναι ιαπωνική λέξη η οποία ετυμολογικά σημαίνει “κύμα στο λιμάνι”. Ουσιαστικά όμως πρόκειται για μία αλληλουχία τεράστιων κυμάτων τα οποία μπορούν να προκληθούν από διάφορες αιτίες, ανάμεσα στις οποίες αρκετά συνήθεις είναι οι σεισμοί. Κατά τη διάρκεια ενός σεισμού ο πυθμένας των ωκεανών ανυψώνεται ή κατακρημνίζεται, μετατοπίζεται ή παίρνει κλίση. Αυτό έχει σαν συνέπεια να δημιουργούνται στο θαλάσσιο επίκεντρο του σεισμού αρχικά μικρά κύματα τα οποία στη συνέχεια κινούνται κυκλικά προς όλες τις κατευθύνσεις με ταχύτητα που μπορεί να φτάσει ακόμα και τα 800 km/h. Το ύψος τους αυξάνεται ραγδαία όσο πλησιάζουν την ακτή και ενώ στον ωκεανό μπορεί να φαίνονται σαν ένα μικρό κύμα, φτάνοντας στην ακτή μπορούν να φτάσουν σε ύψος 40 μέτρων από τη στάθμη της θάλασσας.

1.3.1 Ταχύτητα

Ένα από τα χαρακτηριστικά περιγραφής ενός τσουνάμι είναι η ταχύτητά του. Σύμφωνα με τους σεισμολόγους η ταχύτητα των τσουνάμι υπολογίζεται από τον τύπο $V = \sqrt{g \cdot h}$ όπου g είναι η ένταση της βαρύτητας (ή αλλιώς επιτάχυνση της βαρύτητας) και h το βάθος του νερού στο σημείο που μετρείται η ταχύτητα. Είναι προφανές ότι η ταχύτητα είναι ανάλογη με το βάθος του νερού και στον ανοιχτό ωκεανό προσεγγίζει την ταχύτητα του ήχου.

1.3.2 Ύψος κύματος

Το ύψος του τσουνάμι υπολογίζεται ως η διαφορά του ύψους του κύματος από τον πυθμένα μείον το ύψος του νερού από τον πυθμένα σε φυσιολογικές συνθήκες ή συνθήκες ηρεμίας. Αν δεν είναι γνωστό το ύψος του νερού σε συνθήκες ηρεμίας χρησιμοποιείται ο μέσος όρος της άμπωτης και της πλημμυρίδας του σημείου που καταγράφονται από τους παλιρροιογράφους κάθε περιοχής.

1.3.3 Μέγεθος Τσουνάμι

Το μέγεθος του τσουνάμι υπολογίζεται από τον τύπο που εισήγαγαν οι Imamura και Iida το 1967 $M = \log_2 H_{\max}$ όπου H είναι το μέγιστο καταγεγραμμένο ύψος του τσουνάμι σε μέτρα.

1.3.4 Ένταση Τσουνάμι

Το 1974 οι Soloviev και Go εισήγαγαν έναν τύπο υπολογισμού της έντασης τσουνάμι $I = \log_2 (\sqrt{2} \cdot H)$ όπου H είναι το μέσο καταγεγραμμένο ύψος του τσουνάμι στην ακτή. Παρόλα αυτά το μέγεθος του τσουνάμι και η ένταση είναι μεγέθη που συγχέονται μεταξύ τους και όπως είχε δηλώσει ο Soloviev το 1970 *“If seismological terminology is applied to description of tsunamis, the grades of Imamura-Iida scale must be designated as the intensity of the tsunami and not the magnitude of it... If the seismological terminology is not desired then the term “magnitude” for grades of this scale is quite acceptable”*. Το 2007 ο Δρ. Γεράσιμος Α. Παπαδόπουλος επηρεασμένος από την κλίμακα Mercalli εισήγαγε μια 12-βαθμιαία κλίμακα έντασης του τσουνάμι που υπολογίζει εκ του αποτελέσματος τη σφοδρότητα του κύματος.

Πίνακας 1.2 Κλίμακα έντασης τσουνάμι

Βαθμοί Έντασης Τσουνάμι	
I (<1 m)	Not felt a) Not felt even under the most favourable circumstances. b) No effect. c) No damage.
II (<1 m)	Scarcely felt a) Felt by few people on board in small vessels. Not observed in the coast. b) No effect. c) No damage.
III (<1 m)	Weak a) Felt by most people on board in small vessels. Observed by few people in the coast. b) No effect. c) No damage.
IV (<1 m)	Largely observed a) Felt by all on board in small vessels and by few people on board in large vessels. Observed by most people in the coast. b) Few small vessels move slightly onshore. c) No damage.
V (<1 m)	Strong a) Felt by all on board in large vessels and observed by all in the coast. Few people are frightened and run to higher ground. b) Many small vessels move strongly onshore, few of them crash each other or overturn. Traces of sand layer are left behind in grounds of favourable conditions. Limited flooding of cultivated land.

	c) Limited flooding of outdoors facilities (e.g. gardens) of near-shore structures.
VI (2 m)	Slightly damaging a) Many people are frightened and run to higher ground. b) Most small vessels move violently onshore, or crash strongly each other, or overturn. c) Damage and flooding in a few wooden structures. Most masonry buildings withstand.
VII (4 m)	Damaging a) Most people are frightened and try to run in higher ground. b) Many small vessels damaged. Few large vessels oscillate violently. Objects of variable size and stability overturn and drift. Sand layer and accumulations of pebbles are left behind. Few aquaculture rafts washed away. c) Many wooden structures damaged, few are demolished or washed away. Damage of grade 1 and flooding in a few masonry buildings.
VIII (4 m)	Heavily damaging a) All people escape to higher ground, a few are washed away. b) Most of the small vessels are damaged, many are washed away. Few large vessels are moved ashore or crashed each other. Big objects are drifted away. Erosion and littering in the beach. Extensive flooding. Slight damage in tsunami control forest, stop drifts. Many aquaculture rafts washed away, few partially damaged. c) Most wooden structures are washed away or demolished. Damage of grade 2 in a few masonry buildings. Most RC buildings sustain damage, in a few damage of grade 1 and flooding is observed.
IX (8 m)	Destructive a) Many people are washed away. b) Most small vessels are destructed or washed away. Many large vessels are moved violently ashore, few are destructed. Extensive erosion and littering of the beach. Local ground subsidence. Partial destruction in tsunami control forest, stop drifts. Most aquaculture rafts washed away, many partially damaged. c) Damage of grade 3 in many masonry buildings, few RC buildings suffer from damage grade 2.
X (8 m)	Very destructive a) General panic. Most people are washed away. b) Most large vessels are moved violently ashore, many are destructed or collided with buildings. Small bolders from the sea bottom are moved inland. Cars overturned and drifted. Oil spill, fires start. Extensive ground subsidence. c) Damage of grade 4 in many masonry buildings, few RC buildings suffer from damage grade 3. Artificial embankments collapse, port breakwaters damaged.
XI (16 m)	Devastating b) Lifelines interrupted. Extensive fires. Water backwash drifts cars and other objects in the sea. Big bolders from the sea bottom are moved inland. c) Damage of grade 5 in many masonry buildings. Few RC buildings suffer from damage grade 4, many suffer from damage grade 3.
XII (32 m)	Completely devastating c) Practically all masonry buildings demolished. Most RC buildings suffer from at least damage grade 3.

Πηγή: Papadopoulos G.A. (2007), "Quantification of tsunami: The new 12-point tsunami intensity scale", International Center for Theoretical Physics

Στο επόμενο κεφάλαιο παρουσιάζονται οι βασικές στατιστικές μεθοδολογίες περιγραφής και ανάλυσης του φαινομένου τσουνάμι.

Κεφάλαιο 2.

Στατιστικές Μέθοδοι Ανάλυσης Φαινομένων Τσουνάμι

Σε αυτό το κεφάλαιο παρουσιάζεται όλο το θεωρητικό υπόβαθρο των μεθόδων στατιστικής ανάλυσης που δύναται να χρησιμοποιηθούν για την περιγραφή, πρόβλεψη και ανάλυση των φαινομένων τσουνάμι.

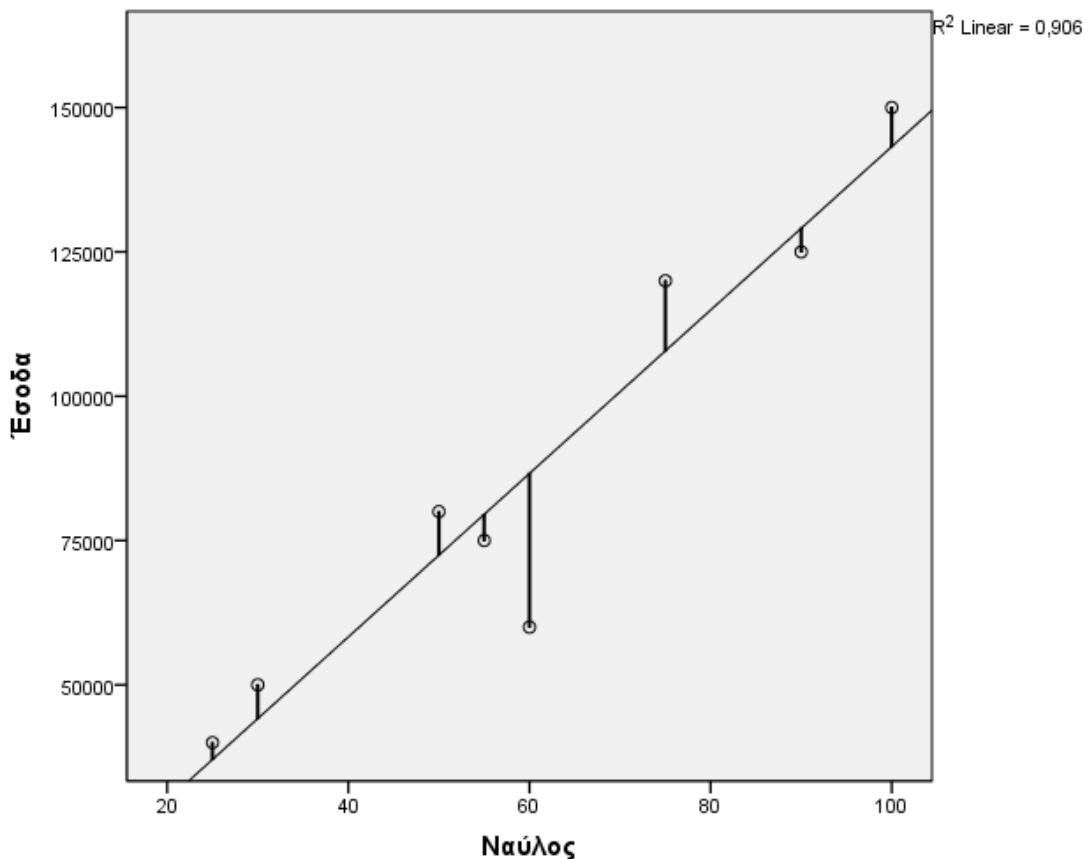
2.1 Ανάλυση Παλινδρόμησης

Για τη διερεύνηση της σχέσης μεταξύ δύο ποσοτικών μεταβλητών ή το πώς μία ή περισσότερες μεταβλητές επηρεάζουν μία άλλη η μεθοδολογία που χρησιμοποιείται ονομάζεται Ανάλυση Παλινδρόμησης.

Αυτή με μεθοδολογία εφαρμόζεται σε οποιοδήποτε πρόβλημα που μπορεί να διακριθούν δύο είδη μεταβλητών, οι εξαρτημένες ή μεταβλητές απόκρισης και οι ανεξάρτητες ή προβλέπουσες μεταβλητές. Ο συνήθης συμβολισμός για τις εξαρτημένες μεταβλητές είναι Y ενώ για τις ανεξάρτητες X . Αν θεωρήσουμε τις μεταβλητές αυτές τυχαίες τότε μπορούμε να εφαρμόσουμε ένα στοχαστικό γραμμικό μοντέλο της μορφής $Y_i = \beta_0 + \beta_1 \cdot X_{i1} + \varepsilon_i$ και να διενεργήσουμε ελέγχους υποθέσεων για τη στατιστική σημαντικότητα των παραμέτρων β_i με βάση την κατανομή που ακολουθούν.

Έτσι για παράδειγμα αν θεωρήσουμε μία τυχαία μεταβλητή Y τα έσοδα μίας ναυτιλιακής εταιρίας και X το κόστος του ναύλου, είναι προφανές ότι τα ποσά αυτά είναι ανάλογα. Αν συλλέξουμε δεδομένα για τις 2 μεταβλητές αυτές μπορούμε να δούμε τη σχέση τους σε ένα σύστημα ορθογώνιων αξόνων. Το γράφημα που θα πάρουμε ονομάζεται διάγραμμα διασποράς και μας δίνει τις πρώτες ενδείξεις για το αν η σχέση των δύο μεταβλητών είναι γραμμική ή όχι. Στη περίπτωση στην οποία η σχέση δεν φαίνεται γραμμική μπορούμε να εφαρμόσουμε κατάλληλους μετασχηματισμούς έτσι ώστε να γίνει γραμμική. Το πρόβλημα τώρα έγκειται στο να φέρουμε την καλύτερη δυνατή ευθεία που διέρχεται πάνω από τα περισσότερα σημεία και αφήνει τις μικρότερες κατακόρυφες αποστάσεις από τα σημεία που είναι εκτός ευθείας όπως φαίνεται στο παρακάτω σχήμα (Εικόνα 1). Η εκτίμηση αυτή δίνεται με τη Μέθοδο Ελαχίστων Τετραγώνων που εκτιμά τις παραμέτρους β_i του παραπάνω στοχαστικού μοντέλου.

Εικόνα 2.1 Διάγραμμα Διασποράς Παράδειγμα



Έτσι προσεγγίζονται τα πραγματικά δεδομένα από ένα στατιστικό μοντέλο στο οποίο μπορούμε να εφαρμόσουμε ελέγχους για την επίλυση του προβλήματος που μελετάται. Για να μπορέσουμε να αξιολογήσουμε την προσέγγιση του μοντέλου χρειάζεται να υποθέσουμε κάποια πράγματα για την κατανομή των υπό μελέτη δεδομένων.

Οι τρεις βασικές προϋποθέσεις της γραμμικής παλινδρόμησης αφορούν τις τυχαίες μεταβλητές ϵ_i . Οι μεταβλητές αυτές ονομάζονται αποκλίσεις ή σφάλματα και ορίζονται ως η διαφορά της πραγματικής τιμής της μεταβλητής απόκρισης από την εκτιμηθείσα από το μοντέλο τιμή. Υποθέτουμε ότι αυτές οι τυχαίες μεταβλητές ακολουθούν κανονική κατανομή $N(0, \sigma^2)$, ότι έχουν κοινή διακύμανση και είναι ανεξάρτητες μεταξύ τους. Υπό αυτές τις προϋποθέσεις μπορούν να χρησιμοποιηθούν έλεγχοι t και F για τη στατιστική σημαντικότητα του μοντέλου και των παραμέτρων β_i .

Στο φαινόμενο τσουνάμι εξαρτημένη μεταβλητή μπορεί να θεωρηθεί το μέγιστο καταγεγραμμένο ύψος του κύματος και ανεξάρτητες το μέγεθος του σεισμού

καθώς και το εστιακό βάθος. Δηλαδή ένα μοντέλο πρόβλεψης για το μέγιστο ύψος του τσουνάμι που μπορεί να εφαρμοστεί είναι της μορφής

$$\text{maximum_height}_i = \beta_0 + \beta_1 \cdot \text{magnitude}_i + \beta_2 \cdot \text{focal_depth}_i + \varepsilon_i$$

2.2 Ανάλυση Bootstrap

Στην περίπτωση που η κατανομή των σφαλμάτων δεν είναι κανονική και δεδομένου ότι η εμφάνιση ενός τσουνάμι κάποια χρονική στιγμή δεν επηρεάζει την εμφάνιση ενός νέου τσουνάμι σε μεταγενέστερη χρονική στιγμή για την εφαρμογή ελέγχων υποθέσεων χρησιμοποιείται η μέθοδος Bootstrap.

Σύμφωνα με τη μέθοδο αυτή διενεργείται τυχαία δειγματοληψία με επανατοποθέτηση από το αρχικό σύνολο δεδομένων, κατασκευάζονται n νέα δείγματα και υπολογίζονται οι στατιστικές συναρτήσεις των ελέγχων με σκοπό την προσομοίωση των ιδιοτήτων των κατανομών των στατιστικών συναρτήσεων με τη χρήση τεχνικών Monte Carlo.

Πιο απλοϊκά, ακόμα και αν δεν υπάρχει κανονικότητα στα δεδομένα υπολογίζεται μια p -value, βασισμένη στην παραπάνω διαδικασία, αρκετά αξιόπιστη για την εξαγωγή συμπερασμάτων για τη στατιστική σημαντικότητα των παραμέτρων της γραμμικής παλινδρόμησης.

2.3 Λογιστική Παλινδρόμηση

Σε ένα πρόβλημα που η μεταβλητή απόκρισης παίρνει δύο τιμές (0 και 1) ενώ οι επεξηγηματικές μεταβλητές είναι ποσοτικές η στατιστική μεθοδολογία που ακολουθείται είναι η εφαρμογή ενός γενικευμένου γραμμικού μοντέλου λογιστικής παλινδρόμησης. Η φύση της μεταβλητής απόκρισης επιτρέπει να θεωρηθεί ότι ακολουθεί κατανομή Bernoulli. Σε ένα τέτοιο πρόβλημα μας ενδιαφέρει να εκτιμήσουμε την πιθανότητα εμφάνισης ενός χαρακτηριστικού με βάση τις επεξηγηματικές μεταβλητές.

Η πιθανότητα εμφάνισης ή όχι ενός τσουνάμι μπορεί να εκτιμηθεί με βάση το μέγεθος του σεισμού και το εστιακό του βάθος. Αν συμβολίσουμε p την πιθανότητα εμφάνισης τσουνάμι το λογιστικό μοντέλο θα έχει τη μορφή

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot \text{magnitude} + \beta_2 \cdot \text{focal_depth}. \quad \text{Ο λόγος } \frac{p}{1-p} \text{ ονομάζεται}$$

σχετική πιθανότητα ενός φαινομένου και μας δίνει το πόσες φορές είναι πιθανότερη η εμφάνιση από τη μη εμφάνιση του φαινομένου. Τα μεγέθη (σε απόλυτη τιμή) των παραμέτρων β_1 και β_2 “δηλώνουν την ταχύτητα με την οποία μεταβάλλεται η πιθανότητα p ” όπως χαρακτηριστικά αναφέρει στις σημειώσεις του μαθήματος “Ανάλυση Διακριτών Δεδομένων” του Μεταπτυχιακού Προγράμματος Σπουδών “Εφαρμοσμένης Στατιστικής” του Πανεπιστημίου Πειραιά ο κ. Γ. Ηλιόπουλος. Για την αξιολόγηση της σημαντικότητας του μοντέλου χρησιμοποιείται ο έλεγχος λόγου πιθανοφανειών που στα στατιστικά πακέτα αναφέρεται ως omnibus test. Ο έλεγχος αυτός συγκρίνει την πιθανοφάνεια του μοντέλου μόνο με το σταθερό όρο με την πιθανοφάνεια του μοντέλου που έχει εφαρμοστεί στα δεδομένα. Η στατιστική σημαντικότητα έγκειται στη μεγάλη διαφορά, αν αυτή υπάρχει, των δύο πιθανοφανειών. Η στατιστική σημαντικότητα των παραμέτρων της λογιστικής παλινδρόμησης ελέγχεται με τον έλεγχο του Wald η στατιστική συνάρτηση του οποίου υπό τη μηδενική υπόθεση ακολουθεί τυπική κανονική κατανομή.

Είναι προφανές ότι η πιθανότητα εμφάνισης τσουνάμι μπορεί να εκτιμηθεί λύνοντας το παραπάνω μοντέλο ως προς p δηλαδή από τη σχέση

$$\hat{p} = \frac{e^{\beta_0 + \beta_1 \cdot magnitude + \beta_2 \cdot focal_depth}}{1 + e^{\beta_0 + \beta_1 \cdot magnitude + \beta_2 \cdot focal_depth}}.$$

2.4 Διατακτική Παλινδρόμηση

Σε ένα πρόβλημα που η μεταβλητή απόκρισης παίρνει παραπάνω από δύο τιμές και η διάταξή τους έχει νόημα, το κλασσικό μοντέλο που χρησιμοποιείται είναι το μοντέλο αθροιστικών logit (cumulative logit model).

Ας υποθέσουμε ότι θέλουμε να μοντελοποιήσουμε την ένταση ενός τσουνάμι με βάση το μέγεθος του σεισμού που το προκάλεσε και το εστιακό βάθος του. Το μοντέλο θα έχει τη μορφή $\log \left[\frac{P(Y \leq j)}{P(Y > j)} \right] = \alpha_j + \beta_1 \cdot magnitude + \beta_2 \cdot focal_depth$.

Είναι προφανές πως ο λογάριθμος των σχετικών πιθανοτήτων για δύο διαφορετικές τιμές του μεγέθους του σεισμού, έστω x_1 και x_2 , και για σταθερό εστιακό βάθος είναι ανάλογος της διαφοράς των τιμών του μεγέθους του σεισμού.

$$\log \left[\frac{P(Y \leq j \mid \text{magnitude} = x_1, \text{focal_depth} = \gamma)}{P(Y > j \mid \text{magnitude} = x_1, \text{focal_depth} = \gamma)} \right] = \alpha_j + \beta_1 \cdot x_1 + \beta_2 \cdot \gamma$$

$$\log \left[\frac{P(Y \leq j \mid \text{magnitude} = x_2, \text{focal_depth} = \gamma)}{P(Y > j \mid \text{magnitude} = x_2, \text{focal_depth} = \gamma)} \right] = \alpha_j + \beta_1 \cdot x_2 + \beta_2 \cdot \gamma$$

$$\log \left[\frac{P(Y \leq j \mid \text{magnitude} = x_2, \text{focal_depth} = \gamma) / P(Y > j \mid \text{magnitude} = x_2, \text{focal_depth} = \gamma)}{P(Y \leq j \mid \text{magnitude} = x_1, \text{focal_depth} = \gamma) / P(Y > j \mid \text{magnitude} = x_1, \text{focal_depth} = \gamma)} \right] = \beta_1 \cdot (x_2 - x_1)$$

Λόγω της παραπάνω ιδιότητας το μοντέλο ονομάζεται και μοντέλο αναλογικών σχετικών πιθανοτήτων (proportional odds). Οι πιθανότητες των κατηγοριών της

έντασης τσουνάμι δίνονται από τον τύπο $\pi_j = \frac{e^{\alpha_j + \beta_1 \cdot \text{magnitude} + \beta_2 \cdot \text{focal_depth}}}{1 + e^{\alpha_j + \beta_1 \cdot \text{magnitude} + \beta_2 \cdot \text{focal_depth}}}$.

Για να μπορέσουμε να ισχυριστούμε ότι το μοντέλο προσαρμόζεται ικανοποιητικά θα πρέπει να γίνει ο έλεγχος παράλληλων γραμμών. Ο έλεγχος αυτός υπό τη μηδενική υπόθεση χρησιμοποιεί το παραπάνω μοντέλο ενώ υπό την

εναλλακτική $\log \left[\frac{P(Y \leq j)}{P(Y > j)} \right] = \alpha_j + \beta_{1j} \cdot \text{magnitude} + \beta_{2j} \cdot \text{focal_depth}$ με μεταβλητά

β για κάθε logit. Διαφορετικά β συνεπάγονται διαφορετικές κλίσεις ευθειών που αναγκαστικά τέμνονται πράγμα που δεν μπορεί να συμβαίνει εφόσον τα $\text{logit}(P(Y \leq j))$ για $j = 1, \dots, J$ είναι διατάξιμα. Επομένως η μη απόρριψη της μηδενικής υπόθεσης στον έλεγχο λόγου πιθανοφανειών είναι ισχυρή ένδειξη καλής προσαρμογής του μοντέλου.

2.5 Discriminant Analysis

Ο W. J. Youden το 1950 εισήγαγε μία συνάρτηση με την οποία μπορούσε να αξιολογήσει διαγνωστικά τεστ για σοβαρές ασθένειες. Η συνάρτηση ορίζεται $J = \text{ευαισθησία} + \text{ειδικότητα} - 1$ και παίρνει τιμές ανάμεσα στο (0,1). Ευαισθησία ορίζεται ως η πιθανότητα το διαγνωστικό τεστ να έδειξε την ύπαρξη μίας ασθένειας δοθέντος ότι ο ασθενής όντως την είχε ενώ ειδικότητα ως η πιθανότητα το διαγνωστικό τεστ να έδειξε ότι ο ασθενής δεν έπασχε από την ασθένεια δοθέντος ότι όντως δεν έπασχε από αυτήν.

Με βάση αυτήν τη συνάρτηση του Youden μπορεί να βρεθεί ένα κάτω άκρο (cut off point ή threshold) σε μία συνεχή μεταβλητή για την πρόβλεψη μίας δίτιμης ονοματικής μεταβλητής. Η μέθοδος που χρησιμοποιείται είναι ο υπολογισμός του δείκτη J του Youden για όλες τις πιθανές τιμές της συνεχούς μεταβλητής, η εύρεση του

μεγίστου J και η αντιστοίχιση της τιμής της συνεχούς μεταβλητής που μεγιστοποίησε το J .

Για την αξιολόγηση της τιμής αυτής χρησιμοποιούνται οι καμπύλες ROC που κατασκευάζονται σε ένα ορθογώνιο σύστημα αξόνων (1- ειδικότητα, ευαισθησία) και υπολογίζεται το εμβαδόν κάτω από την καμπύλη (AUC). Υψηλές τιμές του AUC αποδεικνύουν υψηλή προβλεπτική ικανότητα του threshold ενώ τιμές κοντά στο 0,5 υπονοούν τυχαία ταξινόμηση των δεδομένων. Στα σεισμολογικά δεδομένα το ρόλο της συνεχούς μεταβλητής μπορεί να παίξει το μέγεθος του σεισμού ή ακόμα και η ένταση του.

Εν συνεχεία κατασκευάζεται ένας πίνακας συνάφειας 2×2 όπου στις γραμμές υπάρχει η ένδειξη αν έγινε τσουνάμι ή όχι και στις στήλες πάνω από το συγκεκριμένο μέγεθος σεισμού (ή ένταση) ένδειξη πρόβλεψης τσουνάμι ενώ κάτω ένδειξη αρνητικής πρόβλεψης τσουνάμι. Το άθροισμα των διαγώνιων ποσοστών του πίνακα ορίζεται ως το ποσοστό ορθής ταξινόμησης, υψηλές τιμές του οποίου υποδηλώνουν μεγάλη προβλεπτική ικανότητα του threshold.

2.6 Χωρική Ανάλυση

Οι στατιστικές συναρτήσεις σάρωσης είναι μία μεθοδολογία που ανιχνεύει και αξιολογεί συστάδες σε χωρικό, χρονικό ή χωροχρονικό πλαίσιο. Αυτό συμβαίνει με την επιλογή ενός σχήματος για τη σάρωση ενός χάρτη και την καταγραφή των παρατηρούμενων και των αναμενόμενων περιπτώσεων μέσα στο σχήμα σε κάθε τοποθεσία.

Πιο αναλυτικά ας υποθέσουμε ότι η υπό μελέτη περιοχή, για την εμφάνιση κάποιων γεγονότων, είναι διάστασης $R = [0, L_1] \times [0, L_2]$, με $L_i > 0$ για $i = 1, 2$. Για να μελετήσουμε αν κάποια σημεία της περιοχής εμφανίζουν σημαντικά μεγαλύτερο αριθμών περιπτώσεων θα πρέπει να εφαρμόσουμε ένα πλέγμα πάνω της. Έστω v_1 και v_2 δύο θετικοί ακέραιοι αριθμοί. Σχεδιάζουμε $v_1 + 1$ κάθετες γραμμές (μεσημβρινούς) και $v_2 + 1$ παράλληλες γραμμές χωρίζοντας την αρχική περιοχή σε $v_1 \times v_2$ κελιά. Είναι προφανές πως το μέγεθος κάθε κελιού είναι $h_1 \times h_2$, όπου $h_1 = L_1 / v_1$ και $h_2 = L_2 / v_2$. Έστω μία τυχαία μεταβλητή Y_{ij} , όπου $1 \leq i \leq v_1$ και $1 \leq j \leq v_2$, που αναπαριστά τον αριθμό των γεγονότων τα οποία εμφανίστηκαν στην υπό-περιοχή $[(i-1)h_1, ih_1] \times [(j-1)h_2, jh_2]$. Τέλος ας θεωρήσουμε δύο θετικούς ακέραιους αριθμούς m_1 και m_2 που

καθορίζουν το μέγεθος ενός παραλληλόγραμμου παραθύρου που σαρώνει την περιοχή.
 Η τυχαία μεταβλητή

$$S(i_1, i_2) = \sum_{j=i_2}^{i_2+m_2-1} \sum_{i=i_1}^{i_1+m_1-1} Y_{ij}$$

για $1 \leq i_1 \leq v_1 - m_1 + 1$ και $1 \leq i_2 \leq v_2 - m_2 + 1$ αναπαριστά το συνολικό αριθμό γεγονότων που εμφανίστηκαν στην υπό-περιοχή $[(i_1-1)h_1, (i_1 + m_1 - 1)h_1] \times [(i_2-1)h_2, (i_2 + m_2 - 1)h_2]$. Η στατιστική συνάρτηση σάρωσης που θα χρησιμοποιηθεί είναι ο μέγιστος αριθμός γεγονότων που εμφανίζονται σε περιοχές μεγέθους $(m_1h_1) \times (m_2h_2)$ και ορίζεται

$$S_m = \max\{S(i_1, i_2); 1 \leq i_1 \leq v_1 - m_1 + 1, 1 \leq i_2 \leq v_2 - m_2 + 1\}.$$

Το 1996 οι Chen and Glaz πρότειναν τον έλεγχο για την εύρεση σημαντικών συστάδων στην υπό μελέτη περιοχή,

H_0 : Τυχαία διασπορά των υπό μελέτη γεγονότων στο χώρο
 έναντι της

H_1 : Μη τυχαία διασπορά των υπό μελέτη γεγονότων στο χώρο

με βασική προϋπόθεση την ισονομία και ανεξαρτησία των τυχαίων μεταβλητών Y_{ij} , όπου $1 \leq i \leq v_1$ και $1 \leq j \leq v_2$, που ακολουθούν την κατανομή Poisson(λ_0). Μεγάλες τιμές της S_m δίνουν ενδείξεις για την ύπαρξη τουλάχιστον μίας περιοχής που οι Y_{ij} ακολουθούν κατανομή Poisson(λ_1) με $\lambda_1 > \lambda_0$. Δηλαδή υπό την εναλλακτική υπόθεση ο αναμενόμενος αριθμός γεγονότων είναι στατιστικά μεγαλύτερος ή η εμφάνιση των γεγονότων γίνεται με μεγαλύτερη συχνότητα.

Στη χωρική στατιστική συνάρτηση σάρωσης δημιουργείται ένας κύκλος που σαρώνει το χάρτη, με την ακτίνα του να μεταβάλλεται διαρκώς, και καταγράφει όλους τους κύκλους (ομάδες γειτονικών περιοχών) με τη μεγαλύτερη πιθανοφάνεια. Χρησιμοποιώντας προσομοιώσεις Monte Carlo εφαρμόζονται έλεγχοι υποθέσεων για τη στατιστική σημαντικότητα των συστάδων σε σχέση με όλες τις υπόλοιπες πιθανές. Με τον τρόπο αυτό ομαδοποιούνται περιοχές στις οποίες υπάρχει αυξημένος κίνδυνος εμφάνισης ενός χαρακτηριστικού και εκτιμάται το ποσοστό επιπλέον κινδύνου που εμφανίζει η συγκεκριμένη συστάδα σε σχέση με όλες τις υπόλοιπες με βάση το δεδομένο δείγμα παρατηρήσεων. Στα σεισμολογικά δεδομένα είναι λογικό να χρησιμοποιηθούν οι πρωτεύουσες των χωρών ως συντεταγμένες αντί για τη δημιουργία

κάποιου πλέγματος μέσα στα παράθυρα του οποίου θα εφαρμόζονταν η συγκεκριμένη ανάλυση. Εναλλακτικά μπορούν να χρησιμοποιηθούν τα κέντρα ευρύτερων γεωγραφικών περιοχών αντί των πρωτευουσών.

Κεφάλαιο 3.

Εφαρμογή Στατιστικών Αναλύσεων σε Παγκόσμια Δεδομένα

3.1 Περιγραφή μεταβλητών

Λόγω έλλειψης επαρκών δεδομένων μόνο για τον Ελλαδικό χώρο κατασκευάστηκε ένα παγκόσμιο σετ δεδομένων με στοιχεία που βρέθηκαν στις διαδικτυακές σελίδες <http://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=1&d=1> και http://www.ngdc.noaa.gov/hazard/tsu_db.shtml. Στη μία σελίδα υπάρχουν παγκόσμιες καταγραφές σημαντικότερων σεισμών με ένδειξη αν προκάλεσαν τσουνάμι ή όχι και στην άλλη υπάρχουν παγκόσμιες καταγραφές φαινομένων τσουνάμι με ένδειξη για το αίτιο πρόκλησής τους. Από τα στοιχεία των δύο αυτών σελίδων διασταυρώθηκαν τα φαινόμενα των τσουνάμι με τα στοιχεία των σεισμών που τα προκάλεσαν. Το σετ δεδομένων που κατασκευάστηκε περιέχει τις παρακάτω μεταβλητές:

- ❖ Οι μεταβλητές month, year, day περιγράφουν τις ημερομηνίες κατά τις οποίες καταγράφηκαν οι σεισμοί και τα φαινόμενα τσουνάμι από το 2000 π.Χ. έως το Φεβρουάριο του 2015.
- ❖ flag tsunami: παίρνει την τιμή 1 όταν καταγράφηκε τσουνάμι μετά από σεισμό και την τιμή 0 όταν δεν καταγράφηκε τσουνάμι μετά από σεισμό.
- ❖ focal_depth: είναι το εστιακό βάθος του σεισμού μετρημένο σε χιλιόμετρα
- ❖ eq_primary: είναι το μέγεθος του σεισμού σε οποιαδήποτε κλίμακα μπορούσε αυτό να μετρηθεί
- ❖ eq_intensity: είναι η ένταση του σεισμού μετρημένη στην τροποποιημένη κλίμακα Mercalli
- ❖ country: περιέχει 150 χώρες στις οποίες καταγράφηκαν σεισμοί ή τσουνάμι και υπάρχουν στο σετ δεδομένων
- ❖ eq_location_name: είναι τα ονόματα των περιοχών που καταγράφηκαν σεισμοί
- ❖ latitude: είναι το γεωγραφικό πλάτος μετρημένο σε μοίρες
- ❖ longitude: είναι το γεωγραφικό μήκος μετρημένο σε μοίρες
- ❖ eq_region_code: είναι ευρύτερες γεωγραφικές περιοχές στις οποίες καταγράφηκαν σεισμοί

- ❖ event_validity: είναι η εγκυρότητα της καταγραφής του φαινομένου τσουνάμι. Παίρνει τιμές από το -1 έως το 4, με το -1 να υποδηλώνει λανθασμένη εγγραφή και το 4 σίγουρο τσουνάμι
- ❖ tsu_location_name: είναι τα ονόματα των περιοχών που καταγράφηκαν τσουνάμι
- ❖ tsu_region_code: είναι ευρύτερες γεωγραφικές περιοχές στις οποίες καταγράφηκαν τσουνάμι
- ❖ maximum_water_height: είναι το μέγιστο ύψος των κυμάτων κατά την καταγραφή των φαινομένων τσουνάμι μετρημένο σε μέτρα
- ❖ iida: είναι το μέγεθος των τσουνάμι και δίνεται από τον τύπο $M = \log_2 h$, όπου h είναι το μέγιστο ύψος των κυμάτων κατά την καταγραφή των φαινομένων τσουνάμι
- ❖ soloviev: είναι η ένταση του τσουνάμι όπως αυτή ορίστηκε από τους Soloviev and Go (1974), με τύπο $I = \log_2(2^{1/2} * h)$, όπου h είναι το μέγιστο ύψος των κυμάτων κατά την καταγραφή των τσουνάμι

3.2 Περιγραφική Στατιστική Ανάλυση

Στο σετ δεδομένων που κατασκευάστηκε υπάρχουν 6090 καταγραφές σεισμών εκ των οποίων στις 1987 σημειώθηκαν φαινόμενα τσουνάμι.

Πίνακας 3.1 Συχνότητες σεισμών ανά χώρα

Country			
		Frequency	Percent
Valid	CHINA	573	9,4
	JAPAN	402	6,6
	ITALY	382	6,3
	INDONESIA	366	6,0
	IRAN	362	5,9
	TURKEY	334	5,5
	GREECE	306	5,0
	USA	258	4,2
	PHILIPPINES	204	3,3
	MEXICO	197	3,2
	CHILE	196	3,2
	PERU	175	2,9
	RUSSIA	150	2,5
	TAIWAN	90	1,5
	PAPUA NEW GUINEA	89	1,5
	NEW ZEALAND	88	1,4
	INDIA	87	1,4
	COLOMBIA	74	1,2
	VENEZUELA	62	1,0
	ECUADOR	59	1,0
PORTUGAL	59	1,0	

Στον παραπάνω πίνακα φαίνονται οι χώρες με τη μεγαλύτερη συχνότητα σεισμών στο σετ δεδομένων καθώς και το ποσοστό που αντιστοιχεί στο σύνολο των 6090 καταγραφών. Η Ελλάδα καταλαμβάνει την έβδομη θέση με 306 καταγεγραμμένους σημαντικούς σεισμούς.

Πίνακας 3.2 Συχνότητες σεισμών ανά ευρύτερη γεωγραφική περιοχή

		eq_region_code			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Central, Western and S. Africa	66	1,1	1,1	1,1
	Northern Africa	104	1,7	1,8	2,9
	Antarctica	4	,1	,1	3,0
	East Asia	997	16,4	17,1	20,1
	Central Asia and Caucasus	280	4,6	4,8	24,9
	Kamchatka and Kuril Islands	118	1,9	2,0	26,9
	S. and SE. Asia and Indian Ocean	261	4,3	4,5	31,3
	Atlantic Ocean	12	,2	,2	31,5
	Bering Sea	1	,0	,0	31,6
	Caribbean	154	2,5	2,6	34,2
	Central America	151	2,5	2,6	36,8
	Eastern Europe	51	,8	,9	37,7
	Northern and Western Europe	126	2,1	2,2	39,8
	Southern Europe	816	13,4	14,0	53,8
	Middle East	783	12,9	13,4	67,2
	North America and Hawaii	459	7,5	7,9	75,1
	South America	564	9,3	9,7	84,7
	Central and South Pacific	892	14,6	15,3	100,0
	Total		5839	95,9	100,0
Missing	System	251	4,1		
Total		6090	100,0		

Στον παραπάνω πίνακα φαίνεται η κατανομή συχνοτήτων των σημαντικότερων σεισμών σε ευρύτερες γεωγραφικές περιοχές με την Ανατολική Ασία να καταλαμβάνει τις 997 καταγραφές από το σύνολο των 6090 παρατηρήσεων.

Πίνακας 3.3 Περιγραφικά στοιχεία σεισμών ανά χώρα

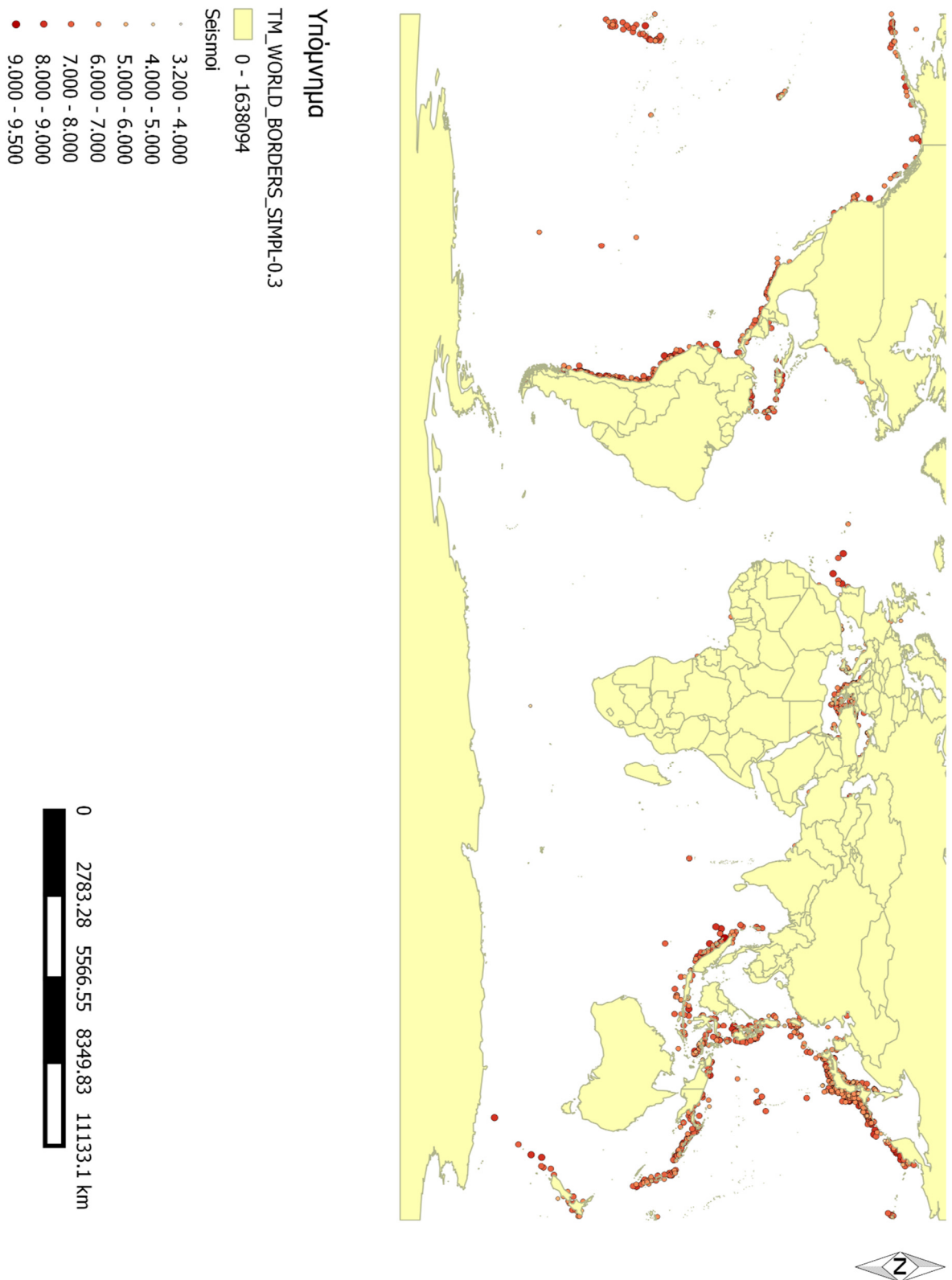
DescriptiveStatistics						
country		N	Minimum	Maximum	Mean	Std. Deviation
CHINA	focal_depth	208	1	460	22,41	32,530
	eq_primary	542	3,40	8,50	6,0306	,85332
	maximum_water_height	7	,75	3,00	1,5071	,84677
	Tsunami magnitude	5	-,30	1,00	,7400	,58138
JAPAN	focal_depth	226	0	516	44,17	59,649
	eq_primary	369	4,60	9,00	6,9829	,72491
	maximum_water_height	186	,02	85,40	3,3922	8,80869
	Tsunami magnitude	196	-3,30	6,40	,4597	1,96311
ITALY	focal_depth	63	1	100	20,16	20,704
	eq_primary	99	3,20	7,90	5,6212	,90489
	maximum_water_height	5	,10	10,00	2,7100	4,15337
INDONESIA	focal_depth	231	5	600	56,94	77,980
	eq_primary	283	4,40	9,10	6,7940	,89506
	maximum_water_height	79	,09	100,00	5,9789	13,25513
	Tsunami magnitude	59	-3,30	4,90	1,3953	1,62696
IRAN	focal_depth	179	3	100	24,16	15,443
	eq_primary	230	4,10	7,90	5,9352	,85643
TURKEY	focal_depth	108	3	80	19,54	16,132
	eq_primary	209	4,00	7,80	6,1469	,89566
	maximum_water_height	4	,53	3,00	1,7375	1,20633
	Tsunami magnitude	2	-,10	1,00	,4500	,77782
GREECE	focal_depth	91	2	130	37,71	33,551
	eq_primary	180	4,00	8,30	6,5167	,72824
	maximum_water_height	20	,10	30,00	3,8475	6,75974
	Tsunami magnitude	4	,30	5,00	1,6750	2,24852
USA	focal_depth	142	0	170	20,04	21,551
	eq_primary	207	1,60	9,20	6,5734	1,20980
	maximum_water_height	48	,03	30,00	3,0804	6,95722
	Tsunami magnitude	26	-3,30	4,30	-,3654	2,60168
PHILIPPINES	focal_depth	91	10	586	44,18	61,864
	eq_primary	116	4,10	8,70	6,9190	,89665
	maximum_water_height	27	,03	9,00	2,0222	2,73399
	Tsunami magnitude	26	-3,30	5,00	,4038	1,96641
MEXICO	focal_depth	81	3	150	46,64	30,258
	eq_primary	112	3,50	8,40	7,0893	,83971
	maximum_water_height	34	,09	11,00	2,9203	3,38063
	Tsunami magnitude	27	-3,30	3,50	,8630	1,78221
CHILE	focal_depth	104	4	180	45,41	29,429
	eq_primary	143	5,00	9,50	7,2839	,78252
	maximum_water_height	52	,09	29,00	4,2644	6,90886
	Tsunami magnitude	44	-3,30	4,60	1,2659	1,98821
PERU	focal_depth	130	1	651	78,00	127,785
	eq_primary	137	4,40	8,80	7,0029	1,07238
	maximum_water_height	23	,38	26,00	4,9196	6,98240
	Tsunami magnitude	22	-1,70	4,60	1,3955	1,76864
RUSSIA	focal_depth	137	3	633	66,82	118,687
	eq_primary	144	3,20	9,00	7,0569	1,00722
	maximum_water_height	64	,03	64,00	4,3313	9,44296
	Tsunami magnitude	51	-3,30	4,90	,3059	2,53522
TAIWAN	focal_depth	44	9	200	38,18	35,930
	eq_primary	88	5,20	8,00	6,5841	,68897
	maximum_water_height	15	,02	10,00	1,4720	3,46473
	Tsunami magnitude	6	-3,30	2,00	-1,8667	2,10016

PAPUA NEW GUINEA	focal_depth	65	0	189	46,32	32,790
	eq_primary	77	4,30	8,00	7,2208	,70791
	maximum_water_height	30	,03	12,00	2,3677	3,07465
	Tsunami magnitude	27	-3,30	3,40	-,0563	2,09781
NEW ZEALAND	focal_depth	64	5	600	39,30	75,306
	eq_primary	79	4,50	8,00	6,7051	,95715
	maximum_water_height	25	,03	9,00	1,3864	2,07840
	Tsunami magnitude	8	-2,18	3,60	,9150	2,17008
INDIA	focal_depth	66	3	141	27,05	23,467
	eq_primary	72	3,50	8,60	5,7667	1,25294
	maximum_water_height	5	,01	1,50	,6520	,68262
	Tsunami magnitude	1	,30	,30	,3000	.
COLOMBIA	focal_depth	40	5	600	64,93	98,520
	eq_primary	50	4,50	8,20	6,6420	,95214
	maximum_water_height	2	1,00	6,00	3,5000	3,53553
	Tsunami magnitude	1	2,30	2,30	2,3000	.
VENEZUELA	focal_depth	24	0	107	19,08	20,656
	eq_primary	29	4,90	8,50	6,4897	1,09033
	maximum_water_height	4	,08	10,00	5,5950	4,20636
	Tsunami magnitude	1	2,60	2,60	2,6000	.
ECUADOR	focal_depth	33	3	207	39,55	45,169
	eq_primary	48	4,00	8,70	6,7688	1,11778
	maximum_water_height	2	,25	5,00	2,6250	3,35876
	Tsunami magnitude	2	,60	1,00	,8000	,28284
PORTUGAL	focal_depth	7	5	33	14,57	9,607
	eq_primary	15	4,60	8,50	6,7533	1,19335
	maximum_water_height	3	,90	18,30	7,2000	9,64209
	Tsunami magnitude	2	1,30	3,60	2,4500	1,62635

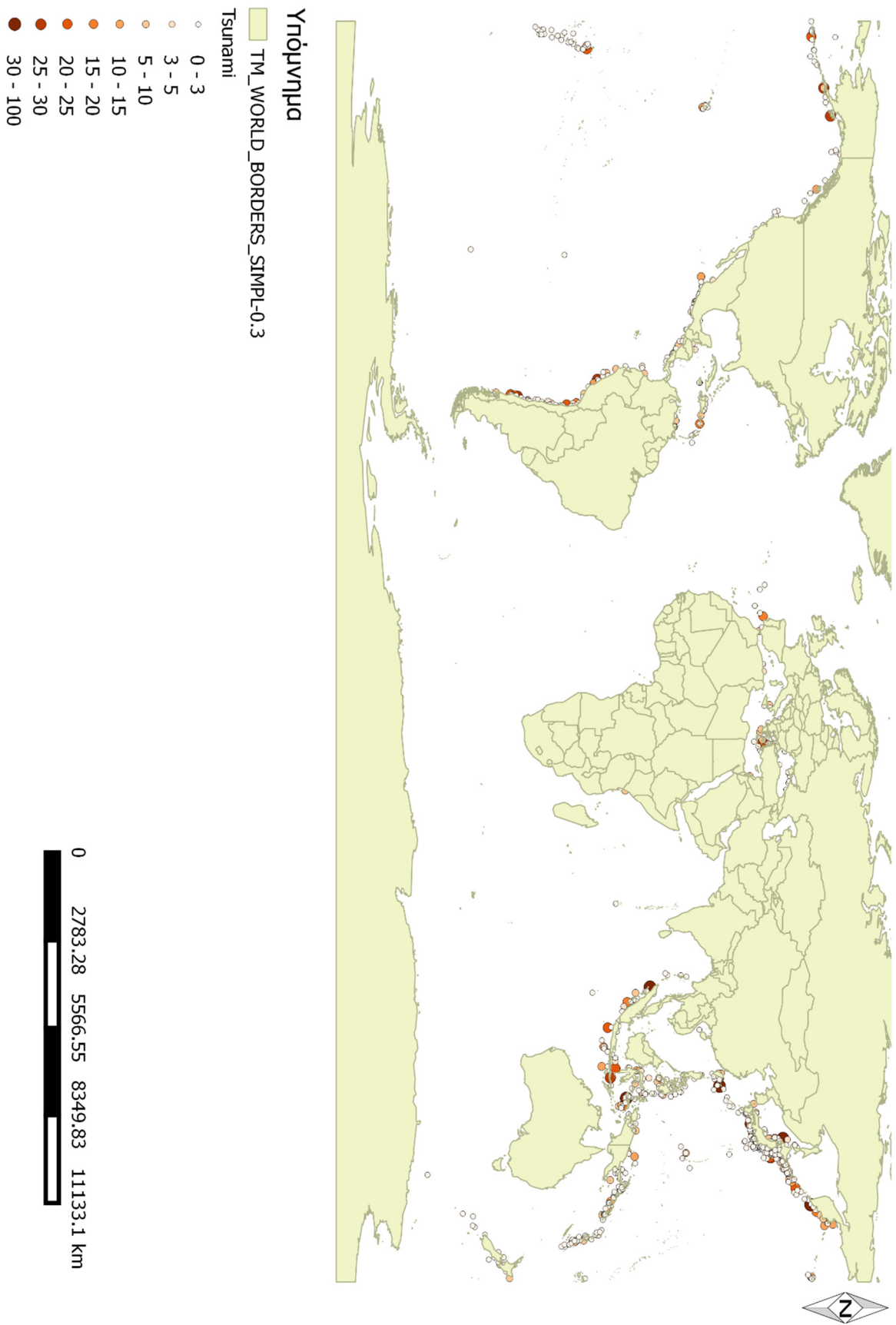
Στον παραπάνω πίνακα φαίνονται τα περιγραφικά μέτρα, των χωρών με τις περισσότερες καταγραφές στο σετ δεδομένων για το μέγεθος του σεισμού, το εστιακό βάθος, το μέγιστο ύψος κύματος σε περιπτώσεις που προκλήθηκε τσουνάμι και το μέγεθος του τσουνάμι. Πιο συγκεκριμένα στην Ελλάδα το ελάχιστο καταγεγραμμένο εστιακό βάθος είναι τα τρία χιλιόμετρα, το μέγιστο τα 130 χιλιόμετρα ενώ το μέσο εστιακό βάθος είναι τα $37,71 \pm 33,551$ χιλιόμετρα. Όσον αφορά το μέγεθος του σεισμού η μέση τιμή είναι $6,5167 \pm 0,73$ ρίχτερ ενώ το μέγιστο κύμα που έχει καταγραφεί είχε ύψος 30 μέτρα.

3.3 Γεωχωρική Ανάλυση

Εικόνα 3.1 Χάρτης σεισμών



Εικόνα 3.2 Χάρτης τσουνάμι



Στους παραπάνω χάρτες φαίνονται τα γεγονότα των σεισμών και των τσουνάμι αντίστοιχα. Αξιοπεριεργό είναι ότι οι σημαντικότεροι σεισμοί συγκεντρώνονται σε παράκτιες περιοχές πράγμα που ίσως εξηγείται από την τοποθεσία των ρηγμάτων κάθε λιθοσφαιρικής πλάκας ή όσων αφορά τουλάχιστον την αμερικανική ήπειρο λόγω των δομών που χρησιμοποιούνται για την κατασκευή κτιρίων. Όπως ήταν αναμενόμενο και η συγκέντρωση των φαινομένων τσουνάμι, βρίσκεται σε παρόμοιες, αν όχι ίδιες, τοποθεσίες.

Όσον αφορά την Ελλάδα υπάρχει πλήθος ισχυρών σεισμών χωρίς όμως να υπάρχουν, τουλάχιστον όσων αφορά την κλίμακα που ορίστηκε, μεγάλα σε ύψος τσουνάμι συγκριτικά με τον υπόλοιπο κόσμο.

3.4 Ανάλυση Παλινδρόμησης

Κατά την αναζήτηση για ένα μοντέλο πρόβλεψης των φαινομένων τσουνάμι εφαρμόστηκε γραμμική παλινδρόμηση σε κάθε χώρα ξεχωριστά καθώς και σε κάθε ευρύτερη γεωγραφική περιοχή που καταγράφηκαν σεισμοί. Στο μοντέλο εξαρτημένη μεταβλητή ορίστηκε το μέγιστο καταγεγραμμένο κύμα ενώ ως ανεξάρτητες το μέγεθος του σεισμού και το εστιακό βάθος.

$$Maximum_water_height_{ij} = \beta_0 + \beta_1 \cdot eq_primary + \beta_2 \cdot focal_depth + \varepsilon_{ij}$$

Όταν οι υποθέσεις της κανονικής γραμμικής παλινδρόμησης δεν τεκμηριώνονταν εφαρμόστηκε η μέθοδος Bootstrap για την εκτίμηση των p-values των συντελεστών της παλινδρόμησης σε κάθε χώρα και σε κάθε ευρύτερη γεωγραφική περιοχή που καταγράφηκαν σεισμοί ξεχωριστά.

Πίνακας 3.4 Παλινδρόμηση ανά γεωγραφική περιοχή

Παλινδρόμηση κατά eq_region						
Region	β_0 (p-value)	β_1 (p-value)	β_2 (p-value)	K-S Test	Levene's Test	Runs Test
East Asia	-21,269 (0,055)	3,37 (0,052)	-0,23 (0,146)	<1%		
Kamchatka and Kuril Islands	-21,942 (0,003)	3,737 (0,003)	-0,049 (0,1)	0,048		
Caribbean	-11,907	1,884	0,09	0,871	0,373	0,627

	(0,11)	(0,07)	(0,185)			
North America and Hawaii	-12,948 (0,097)	1,957 (0,081)	0,024 (0,374)	<1‰		
South America	-49,394 (<1‰)	6,857 (<1‰)	0,039 (0,242)	0,276	0,006	0,202

Στον παραπάνω πίνακα στην πρώτη στήλη φαίνεται η τιμή και η p-value του ελέγχου στατιστικής σημαντικότητας για τον σταθερό όρο, στη δεύτερη στήλη φαίνεται η τιμή και η p-value του ελέγχου στατιστικής σημαντικότητας για την παράμετρο β_1 του μεγέθους του σεισμού, στην τρίτη στήλη φαίνεται η τιμή και η p-value του ελέγχου στατιστικής σημαντικότητας για την παράμετρο β_2 του εστιακού βάθους του σεισμού, στην τέταρτη στήλη φαίνεται η p-value του ελέγχου κανονικότητας (K-S test), στην πέμπτη στήλη φαίνεται η p-value του ελέγχου ομοσκεδαστικότητας (Levene's test) και στην τελευταία στήλη φαίνεται η p-value του ελέγχου ανεξαρτησίας (Runs test).

Ο σταθερός όρος της παλινδρόμησης συνήθως δεν έχει φυσική ερμηνεία παρόλα αυτά είναι η αναμενόμενη τιμή της εξαρτημένης μεταβλητής όταν όλες οι επεξηγηματικές πάρουν την τιμή 0. Η ερμηνεία του β_1 είναι η αναμενόμενη αύξηση ή μείωση του μέγιστου ύψους κύματος ενός τσουνάμι εάν αυξηθεί κατά μία μονάδα το μέγεθος του σεισμού ενώ το εστιακό βάθος παραμένει σταθερό. Τέλος το β_2 είναι η αναμενόμενη αύξηση ή μείωση του μέγιστου ύψους κύματος ενός τσουνάμι εάν αυξηθεί κατά μία μονάδα το εστιακό βάθος του σεισμού ενώ το μέγεθος παραμένει σταθερό.

Όπως φαίνεται στον παραπάνω πίνακα μόνο στη γεωγραφική περιοχή της Καραϊβικής ισχύουν όλες οι προϋποθέσεις της γραμμικής παλινδρόμησης και το εστιακό βάθος κρίνεται η μόνη στατιστικά σημαντική μεταβλητή που πρέπει να μπει στο μοντέλο πρόβλεψης με p-value ίση με 7%. Για όλες τις υπόλοιπες περιοχές εφαρμόστηκε μέθοδος bootstrap για την εκτίμηση των p-values των συντελεστών του μοντέλου και σε όλες υπάρχουν ενδείξεις αλλού περισσότερο ισχυρές και αλλού λιγότερο για τη στατιστική σημαντικότητα του μεγέθους του σεισμού. Αντίθετα το εστιακό βάθος δεν φαίνεται να είναι στατιστικά σημαντικό.

Πίνακας 3.5 Παλινδρόμηση ανά χώρα

Παλινδρόμηση κατά country						
Region	β_0 (p-value)	β_1 (p-value)	β_2 (p-value)	K-S Test	Levene's Test	Runs Test
Chile	-61,244 ($<1\%$)	8,985 ($<1\%$)	-0,108 (0,176)	0,648	0,386	0,193
Greece	-56,302 (0,07)	8,673 (0,05)	0,068 (0,718)	0,789	0,224	0,364
Japan	-27,608 (0,041)	4,295 (0,037)	-0,023 (0,198)	$<1\%$		
Philippines	-9,332 (0,128)	1,684 (0,036)	-0,02 (0,568)	0,369	0,001	0,531
Russia	-19,224 (0,001)	3,405 ($<1\%$)	-0,056 (0,107)	0,053	0,002	0,427
Solomon Islands	-20,111 (0,025)	3,055 (0,014)	0,001 (0,981)	0,471	0,086	1
Tonga	-13,856 (0,05)	1,952 (0,039)	-0,007 (0,236)	0,933	0,936	0,495

Η ίδια ανάλυση εφαρμόστηκε και κατά χώρα και στον πίνακα 3.5 παρουσιάζονται τα αποτελέσματά της. Στη Χιλή, την Ελλάδα και την Τόγκα οι προϋποθέσεις του μοντέλου γραμμικής παλινδρόμησης ισχύουν όλες σε αντίθεση με τις υπόλοιπες χώρες. Και σε αυτόν τον πίνακα δεν υπάρχουν ενδείξεις για τη στατιστική σημαντικότητα του εστιακού βάθους. Το μοντέλο για την Ελλάδα γράφεται $Maximum_water_height_{ij} = -56,302 + 8,673 \cdot eq_primary$.

3.5 Λογιστική Παλινδρόμηση

Για να είναι εφικτή η εκτίμηση της πιθανότητας να εμφανιστεί τσουνάμι μετά από σεισμό εφαρμόστηκε μοντέλο λογιστικής παλινδρόμησης με εξαρτημένη μεταβλητή τη *flag_tsunami* και ανεξάρτητες το μέγεθος του σεισμού και το εστιακό βάθος.

$$\logit(flag_tsunami) = \beta_0 + \beta_1 \cdot eq_primary + \beta_2 \cdot focal_depth$$

Πίνακας 3.6 Λογιστική παλινδρόμηση ανά γεωγραφική περιοχή

Λογιστική Παλινδρόμηση by region										
Region	N		Deviance		Omnibus Test		B			Hypothesis test
	Tsunami not occurred	Tsunami occurred	value	df	LR	p-value	Intercept Exp(b ₀)	eq primary Exp(b ₁)	focal depth Exp(b ₂)	p-value
Northern Africa	28	7	22,494	29	12,534	0,002	-13,290 (1,692*10 ⁻⁶)	1,959 (7,095)	0,013 (1,013)	0,003 (b ₀) 0,007 (b ₁) 0,805 (b ₂)
East Asia	249	168	359,225	317	119,342	<1%	-9,241 (9,7*10 ⁻⁵)	1,423 (4,149)	-0,015 (0,977)	<1% (b ₀ ,b ₁ ,b ₂)
Kamchatka and Kuril Islands	40	70	104,819	97	26,204	<1%	5,456 (234,144)	-0,530 (0,589)	-0,017 (0,983)	0,035 (b ₀) 0,117 (b ₁) 0,008 (b ₂)
S. and SE. Asia and Indian Ocean	158	28	93,534	161	61,287	<1%	-14,189 (6,886*10 ⁻⁷)	1,856 (6,339)	-0,001 (0,999)	<1% (b ₀ ,b ₁) 0,834 (b ₂)
Caribbean	21	28	59,316	46	7,609	0,022	-4,522 (0,011)	0,785 (2,193)	-0,013 (0,987)	0,060 (b ₀) 0,027 (b ₁) 0,158 (b ₂)
Central America	67	15	49,394	73	25,865	<1%	-16,701 (5,585*10 ⁻⁸)	2,245 (9,444)	-0,002 (0,998)	<1% (b ₀ ,b ₁) 0,877 (b ₂)
Southern Europe	175	45	167,473	175	29,082	<1%	-8,438 (2,164*10 ⁻⁴)	1,208 (3,346)	-0,008 (0,992)	<1% (b ₀ ,b ₁) 0,280 (b ₂)
Middle East	285	8	48,687	224	18,112	<1%	-14,669 (4,26*10 ⁻⁷)	1,657 (5,241)	0,019 (1,019)	<1% (b ₀) 0,001 (b ₁) 0,346 (b ₂)
North America and Hawaii	144	85	230,538	204	59,413	<1%	-8,686 (1,689*10 ⁻⁴)	1,276 (3,581)	-0,018 (0,982)	<1% (b ₀ ,b ₁) 0,007 (b ₂)
South America	213	116	304,918	272	85,193	<1%	-8,867 (1,409*10 ⁻⁴)	1,234 (3,436)	-0,007 (0,993)	<1% (b ₀ ,b ₁) 0,006 (b ₂)
Central and South Pacific	311	276	531,14	399	66,223	<1%	-5,275 (0,005)	0,778 (2,176)	-0,006 (0,994)	<1% (b ₀ ,b ₁ ,b ₂)

Στον παραπάνω πίνακα παρουσιάζονται, κατά ευρύτερη γεωγραφική περιοχή:

- Στην πρώτη στήλη ο αριθμός των περιπτώσεων σεισμών που δεν εμφανίστηκε τσουνάμι
- Στη δεύτερη στήλη ο αριθμός των περιπτώσεων των σεισμών που εμφανίστηκαν τσουνάμι
- Στην τρίτη στήλη η τιμή της απόκλισης του μοντέλου που προσαρμόσαμε. Η απόκλιση ορίζεται ως -2 φορές τη διαφορά του λογαρίθμου της πιθανοφάνειας του προσαρμοσμένου μοντέλου από το λογάριθμο της πιθανοφάνειας του κορεσμένου μοντέλου (μοντέλο που έχει μία μεταβλητή για κάθε παρατήρηση). Διαισθητικά, όσο πιο μικρή είναι η απόκλιση ενός μοντέλου, τόσο πιο κοντά είναι στο κορεσμένο μοντέλο, και αυτό παρέχει ένδειξη καλής προσαρμογής
- Στην τέταρτη στήλη η διαφορά των βαθμών ελευθερίας του προσαρμοσμένου μοντέλου από τους βαθμούς ελευθερίας του κορεσμένου

- Στην πέμπτη στήλη ο λόγος της πιθανοφάνειας του προσαρμοσμένου μοντέλου προς την πιθανοφάνεια του μοντέλου μόνο με το σταθερό όρο
- Στην έκτη στήλη η p-value του Omnibus test
- Στην έβδομη στήλη η τιμή του σταθερού όρου καθώς και η εκθετική τιμή του. Η ερμηνεία του σταθερού όρου είναι: ο λογάριθμος της σχετική πιθανότητας εμφάνισης τσουνάμι είναι β_0 φορές το λογάριθμο της σχετικής πιθανότητας μη εμφάνισης τσουνάμι όταν το μέγεθος του σεισμού και το εστιακό βάθος είναι 0, ουσιαστικά όταν δεν έχει γίνει σεισμός
- Στην όγδοη στήλη η τιμή του β_1 καθώς και η εκθετική τιμή του. Η ερμηνεία του β_1 είναι: ο λογάριθμος της σχετική πιθανότητας εμφάνισης τσουνάμι είναι β_1 φορές το λογάριθμο της σχετικής πιθανότητας μη εμφάνισης τσουνάμι όταν το μέγεθος του σεισμού αυξηθεί κατά μία μονάδα και το εστιακό βάθος παραμένει σταθερό
- Στην ένατη στήλη η τιμή του β_2 καθώς και η εκθετική τιμή του. Η ερμηνεία του β_2 είναι: ο λογάριθμος της σχετική πιθανότητας εμφάνισης τσουνάμι είναι β_2 φορές το λογάριθμο της σχετικής πιθανότητας μη εμφάνισης τσουνάμι όταν το εστιακό βάθος του σεισμού αυξηθεί κατά μία μονάδα και το μέγεθος παραμένει σταθερό
- Στην δέκατη στήλη οι p-values του ελέγχου Wald και των τριών συντελεστών της λογιστικής παλινδρόμησης.

Πιο συγκεκριμένα στη γεωγραφική περιοχή της Βόρειας Αφρικής καταγράφηκαν 35 σεισμοί εκ των οποίων στους 7 εμφανίστηκε τσουνάμι. Η p-value του Omnibus test υπολογίστηκε 2%, επομένως το μοντέλο κρίνεται στατιστικά σημαντικό έναντι του μοντέλου μόνο με το σταθερό όρο. Στον έλεγχο για τη στατιστική σημαντικότητα των συντελεστών της λογιστικής παλινδρόμησης σημαντικοί κρίθηκαν ο σταθερός όρος και το μέγεθος του σεισμού. Δηλαδή το μοντέλο για τη Βόρεια Αφρική είναι

$$\log it(flag_tsunami) = -13.29 + 1.959 \cdot eq_primary$$

Η αύξηση κατά μία μονάδα στο μέγεθος του σεισμού επιφέρει αύξηση 1,959 στο λογάριθμο της σχετικής πιθανότητας εμφάνισης τσουνάμι για την ευρύτερη γεωγραφική περιοχή της Βόρειας Αφρικής.

Πίνακας 3.7 Λογιστική παλινδρόμηση ανά χώρα

Λογιστική Παλινδρόμηση by country										
Country	N		Deviance		Omnibus Test		B			Hypothesis test
	Tsunami not occurred	Tsunami occurred	value	df	LR	p-value	Intercept Exp(b ₀)	eq primary Exp(b ₁)	focal depth Exp(b ₂)	p-value
CHILE	31	71	103,795	86	18,719	<1%	-5,242 (0,005)	1,040 (2,830)	-0,026 (0,974)	0,029 (b ₀) 0,003 (b ₁) 0,002 (b ₂)
CHINA	202	5	38,921	158	5,417	0,067	-9,192 (1,019*10 ⁻⁴)	1,037 (2,821)	-0,058 (0,849)	0,008 (b ₀) 0,048 (b ₁) 0,281 (b ₂)
COSTA RICA	10	6	8,662	13	12,508	0,002	-27,273 (1,431*10 ⁻¹²)	3,873 (48,066)	0,025 (1,026)	0,054 (b ₀) 0,063 (b ₁) 0,655 (b ₂)
ECUADOR	29	3	12,098	28	7,815	0,2	-20,664 (1,061*10 ⁻⁹)	2,470 (11,824)	0,009 (1,010)	0,044 (b ₀) 0,061 (b ₁) 0,425 (b ₂)
GREECE	58	28	89,054	78	13,934	0,001	-9,005 (1,228*10 ⁻⁴)	1,388 (4,008)	-0,018 (0,964)	<1% (b ₀) 0,001 (b ₁) 0,063 (b ₂)
GUATEMALA	20	5	9,944	21	12,303	0,002	-29,886 (1,048*10 ⁻¹³)	4,160 (64,094)	-0,028 (0,973)	0,056 (b ₀) 0,055 (b ₁) 0,315 (b ₂)
INDIA	58	7	17,811	55	26.606	<1%	-19,769 (2,597*10 ⁻⁹)	2,753 (15,687)	-0,028 (0,972)	0,005 (b ₀ , b ₁) 0,298 (b ₂)
INDONESIA	137	94	231,355	200	37,146	<1%	-7,094 (0,001)	0,996 (2,709)	-0,002 (0,998)	<1% (b ₀ , b ₁) 0,308 (b ₂)
JAPAN	73	153	205,554	188	47,262	<1%	-5,152 (0,006)	1,015 (2,759)	-0,024 (0,976)	0,001 (b ₀) <1% (b ₁ , b ₂)
MEXICO	54	27	76,128	71	18,669	<1%	-11,298 (1,24*10 ⁻⁵)	1,605 (4,977)	-0,023 (0,978)	0,001 (b ₀ , b ₁) 0,050 (b ₂)
NEW ZEALAND	35	29	52,652	51	17,826	<1%	-9,222 (9,889*10 ⁻⁷)	1,421 (4,140)	-0,022 (0,979)	0,001 (b ₀ , b ₁) 0,125 (b ₂)
PAKISTAN	34	4	13,002	33	12,572	0,002	-17,061 (3,893*10 ⁻⁸)	2,526 (12,508)	-0,069 (0,933)	0,020 (b ₀) 0,022 (b ₁) 0,230 (b ₂)
PAPUA NEW GUINEA	26	39	62,907	50	7,041	0,030	-1,579 (0,206)	0,434 (1,544)	-0,025 (0,975)	0,564 (b ₀) 0,271 (b ₁) 0,038 (b ₂)
PERU	86	37	110,822	111	34,075	<1%	-9,812 (5,478*10 ⁻⁵)	1,287 (3,621)	-0,006 (0,987)	<1% (b ₀ , b ₁) 0,104 (b ₂)
PHILIPPINES	46	45	69,689	73	10,724	0,005	-5,219 (0,005)	0,773 (2,165)	-0,004 (0,996)	0,004 (b ₀) 0,003 (b ₁) 0,390 (b ₂)
RUSSIA	64	73	152,463	120	15,228	<1%	-1,626 (0,197)	0,319 (1,375)	-0,009 (0,991)	0,215 (b ₀) 0,088 (b ₁) 0,018 (b ₂)
TAIWAN	31	13	40,924	38	12,489	0,002	-13,030 (2,192*10 ⁻⁶)	2,036 (7,663)	-0,058 (0,944)	0,009 (b ₀ , b ₁) 0,062 (b ₂)
TURKEY	103	5	26,369	93	14,123	0,001	-16,993 (4,169*10 ⁻⁸)	2,008 (7,446)	0,036 (1,037)	0,001 (b ₀) 0,007 (b ₁) 0,232 (b ₂)
USA	87	54	137,190	128	44,937	<1%	-8,920 (<1%)	1,297 (3,660)	-0,005 (0,995)	<1% (b ₀ , b ₁) 0,605 (b ₂)
VENEZUELA	15	8	17,271	20	12,449	0,002	-14,149 (7,164*10 ⁻⁷)	2,007 (7,442)	0,029 (1,030)	0,017 (b ₀) 0,026 (b ₁) 0,278 (b ₂)

Στον παραπάνω πίνακα παρουσιάζονται αντίστοιχα τα αποτελέσματα για κάθε χώρα ξεχωριστά. Πιο συγκεκριμένα, για την Ελλάδα, καταγράφηκαν 86 σεισμοί εκ των οποίων στους 28 εμφανίστηκαν τσουνάμι. Η p-value του Omnibus test υπολογίστηκε στο 1%, επομένως το μοντέλο κρίνεται στατιστικά σημαντικό έναντι του μοντέλου μόνο με το σταθερό όρο. Στον έλεγχο στατιστικής σημαντικότητας των συντελεστών της λογιστικής παλινδρόμησης ο σταθερός όρος και το μέγεθος του σεισμού κρίνονται στατιστικά σημαντικοί ενώ το εστιακό βάθος δίνει ενδείξεις για σημαντικότητα σε επίπεδο $\alpha=7\%$.

$$\text{logit}(\text{flag_tsunami}) = -9.005 + 1.388 \cdot \text{eq_primary} - 0.018 \cdot \text{focal_depth}$$

Η αύξηση κατά μία μονάδα στο μέγεθος του σεισμού επιφέρει αύξηση 1,388 στο λογάριθμο της σχετικής πιθανότητας εμφάνισης τσουνάμι για την Ελλάδα με δεδομένο το εστιακό βάθος σεισμού. Αντίστοιχα η αύξηση κατά μία μονάδα στο εστιακό βάθος του σεισμού επιφέρει μείωση στο λογάριθμο της σχετικής πιθανότητας εμφάνισης τσουνάμι κατά 0,018 για δεδομένο μέγεθος σεισμού.

3.6 Διατακτική Παλινδρόμηση

Με σκοπό την πρόβλεψη της σφοδρότητας του φαινομένου τσουνάμι εφαρμόστηκε ένα μοντέλο διατακτικής παλινδρόμησης με επεξηγηματικές μεταβλητές το μέγεθος του σεισμού και το εστιακό βάθος. Εξαρτημένη μεταβλητή ορίστηκε η ένταση του τσουνάμι όπως αυτή ορίστηκε από τον δρ. Γεράσιμο Παπαδόπουλο εφόσον η ένταση του Soloviev είναι συνεχής μεταβλητή και δε θα ίσχυαν οι προϋποθέσεις του μοντέλου εάν εφαρμοζόταν.

$$\log \left[\frac{P(TI_GP \leq j)}{P(TI_GP > j)} \right] = \alpha_j + \beta_1 \cdot \text{eq_primary} + \beta_2 \cdot \text{focal_depth}$$

Πίνακας 3.8 Πίνακας συχνοτήτων επιπέδων έντασης Τσουνάμι

Case Processing Summary			
		N	Marginal Percentage
TI_GP	1-5	376	65,5%
	6	60	10,5%
	7-8	45	7,8%
	9-10	42	7,3%
	11	33	5,7%
	12	18	3,1%
Valid		574	100,0%
Missing		5516	
Total		6090	

Στον παραπάνω πίνακα παρατηρούμε πως το 65,5% των 574 καταγραφών τσουνάμι είχαν ένταση 1 έως 5 βαθμών.

Πίνακας 3.9 Omnibus test διατακτικής παλινδρόμησης

Model Fitting Information				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	1197,114			
Final	1043,103	154,011	2	,000

Link function: Logit.

Το μοντέλο κρίνεται στατιστικά σημαντικό σύμφωνα με το Omnibus test η p-value του οποίου είναι μικρότερη του 1%, δηλαδή το μοντέλο που κατασκευάστηκε διαφέρει στατιστικά σημαντικά από το μοντέλο μόνο με το σταθερό όρο.

Πίνακας 3.10 Μέτρα καλής προσαρμογής

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	1911,763	1973	,835
Deviance	935,225	1973	1,000

Link function: Logit.

Υπάρχουν ενδείξεις καλής προσαρμογής του μοντέλου στα δεδομένα εφόσον οι p-values των μέτρων καλής προσαρμογής είναι πάρα πολύ υψηλές.

Πίνακας 3.11 Δείκτες ψευδο-R²

Pseudo R-Square	
Cox and Snell	,235
Nagelkerke	,260
McFadden	,114
Link function: Logit.	

Σύμφωνα με τους παραπάνω δείκτες το μοντέλο που εφαρμόσαμε ερμηνεύει από 11,4% έως 26% της μεταβλητότητας του συνόλου των δεδομένων.

Πίνακας 3.12 Πίνακας συντελεστών διατακτικής παλινδρόμησης

Parameter Estimates								
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[TI_GP = 1]	14,493	1,291	126,028	1	,000	11,963	17,023
	[TI_GP = 2]	15,126	1,305	134,336	1	,000	12,568	17,684
	[TI_GP = 3]	15,752	1,319	142,683	1	,000	13,168	18,337
	[TI_GP = 4]	16,583	1,336	154,017	1	,000	13,964	19,202
	[TI_GP = 5]	17,815	1,365	170,382	1	,000	15,140	20,489
Location	eq_primary	1,892	,173	119,564	1	,000	1,553	2,231
	focal_depth	,000	,002	,047	1	,828	-,005	,004
Link function: Logit.								

Στον παραπάνω πίνακα στην πρώτη στήλη είναι οι εκτιμήσεις των παραμέτρων α_j καθώς και των β_1 και β_2 . Στη δεύτερη στήλη είναι τα τυπικά σφάλματα των εκτιμήσεων των παραμέτρων της παλινδρόμησης, ακολουθούν οι τιμές του ελέγχου του Wald για τη σημαντικότητα των συντελεστών με τους βαθμούς ελευθερίας και οι p-values των ελέγχων με ένα 95% διάστημα εμπιστοσύνης.

Στο παραπάνω μοντέλο ο λογάριθμος της σχετικής πιθανότητας η ένταση του τσουνάμι να είναι μεταξύ 1 και 5 βαθμών είναι 14,5 περίπου φορές το λογάριθμο της σχετικής πιθανότητας η ένταση του τσουνάμι να είναι 12 βαθμών όταν το μέγεθος και το εστιακό βάθος του σεισμού έχουν πάρει την τιμή 0. Επίσης ο λογάριθμος της σχετικής πιθανότητας η ένταση του τσουνάμι να είναι μεταξύ 1 και 5 βαθμών είναι 1,89 περίπου φορές το λογάριθμο της σχετικής πιθανότητας η ένταση του τσουνάμι να είναι 12 βαθμών όταν το μέγεθος του σεισμού αυξηθεί κατά μία μονάδα και το εστιακό βάθος παραμείνει σταθερό.

Το μέγεθος του σεισμού κρίνεται στατιστικά σημαντικό με p-value μικρότερη του 1% ενώ το εστιακό βάθος μη σημαντικό με p-value ίση με 0,828.

Για άλλη μια φορά το μέγεθος του σεισμού είναι το μόνο σημαντικό με βάση το σύνολο των δεδομένων για στην εφαρμογή του μοντέλου.

Πίνακας 3.13 Έλεγχος παράλληλων γραμμών

Test of Parallel Lines^c				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	1043,103			
General	1037,881 ^a	5,222 ^b	8	,734
The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.				

Στον έλεγχο παράλληλων γραμμών δεν υπάρχουν στατιστικές ενδείξεις για την απόρριψη της μηδενικής υπόθεσης επομένως το μοντέλο προσαρμόζεται καλά.

3.7 Discriminant Analysis

Με σκοπό την εύρεση ενός κάτω άκρου, για το μέγεθος του σεισμού ή την ένταση του σεισμού, ικανό να προκαλέσει τσουνάμι, και κατ' επέκταση την πρόβλεψη τέτοιων καταστροφικών φαινομένων, πραγματοποιήθηκε ανάλυση των δεδομένων με καμπύλες Roc, βρέθηκε το μέγιστο της συνάρτησης του Youden και εν συνεχεία αξιολογήθηκε το αποτέλεσμα από το ποσοστό ορθής ταξινόμησης του συγκεκριμένου μεγέθους σεισμού.

Πίνακας 3.14 AUC ανά γεωγραφική περιοχή – Μέγεθος σεισμού

Area Under the Curve					
Test Result Variable(s):eq_primary					
eq_region_code	Area	Std. Error	AsymptoticSig.	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
Northern Africa	,889	,048	,000	,796	,982
East Asia	,779	,015	,000	,749	,809
Kamchatka and Kuril Islands	,337	,052	,004	,235	,440
S. and SE. Asia and Indian Ocean	,857	,027	,000	,805	,910
Caribbean	,679	,074	,017	,535	,824
Central America	,857	,038	,000	,783	,932
Southern Europe	,737	,029	,000	,680	,794
Middle East	,791	,035	,000	,723	,859
North America and Hawaii	,681	,029	,000	,624	,739
South America	,704	,025	,000	,654	,754
Central and South Pacific	,586	,022	,000	,543	,628
a. Under the nonparametric assumption					
b. Null hypothesis: true area = 0.5					

Στον παραπάνω πίνακα φαίνονται όλες οι ευρύτερες γεωγραφικές περιοχές στις οποίες το AUC είναι στατιστικά σημαντικό με βάση το μέγεθος του σεισμού.

Πίνακας 3.15 Ποσοστά ορθής ταξινόμησης ανά γεωγραφική περιοχή – Μέγεθος σεισμού

Discriminant Analysis			
Region	Youden's Maximum	Magnitude(c)	Ποσοστό Ορθής Ταξινόμησης (%)
Northern Africa	0,74	5,65	78,9
East Asia	0,5	6,35	71,2
Kamchatka and Kuril Islands	0,14	8,05	44,1
S. and SE. Asia and Indian Ocean	0,59	6,15	70,5
Caribbean	0,36	6,15	64,3
Central America	0,67	6,75	86,1
Southern Europe	0,45	6,45	82,7
Middle East	0,56	6,35	79,6
North America and Hawaii	0,3	6,25	62,5
South America	0,33	6,85	62,8
Central and South Pacific	0,19	6,45	58,1

Στον παραπάνω πίνακα φαίνονται ποιο μέγεθος σεισμού μεγιστοποιεί τη συνάρτηση του Youden καθώς επίσης και το ποσοστό ορθής ταξινόμησης για κάθε γεωγραφική περιοχή ξεχωριστά. Πιο αναλυτικά για τη Νότια Ευρώπη βρέθηκε ότι ένα μέγεθος σεισμού της τάξης των 6,75 ρίχτερ είναι ικανό να προκαλέσει τσουνάμι. Το ποσοστό ορθής ταξινόμησης του συγκεκριμένου μεγέθους σεισμού είναι 86,1%, πράγμα που σημαίνει ότι στα δεδομένα που έχουν καταγραφεί αν χρησιμοποιούνταν το συγκεκριμένο μέγεθος σεισμού για την πρόβλεψη εμφάνισης τσουνάμι στο 86,1% των περιπτώσεων η πρόβλεψη θα ήταν σωστή ενώ στο 13,9% των περιπτώσεων λανθασμένη.

Αντίστοιχη ανάλυση πραγματοποιήθηκε και για κάθε χώρα ξεχωριστά με τα σημαντικότερα αποτελέσματα να παρατίθενται στους δύο πίνακες που ακολουθούν.

Πίνακας 3.16 AUC ανά χώρα – Μέγεθος σεισμού

Area Under the Curve					
Test Result Variable(s):eq_primary					
country	Area	Std. Error	AsymptoticSig.	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
ALBANIA	,793	,080	,005	,637	,950
ALGERIA	,919	,053	,003	,815	1,000
AUSTRALIA	,949	,048	,006	,854	1,000
CHILE	,598	,052	,067	,496	,701
CHINA	,610	,074	,081	,464	,755
COSTA RICA	,950	,051	,000	,851	1,000
CROATIA	1,000	,000	,046	1,000	1,000
CUBA	,000	,000	,083	,000	,000
ECUADOR	,811	,137	,074	,542	1,000
EGYPT	,917	,115	,077	,691	1,000
EL SALVADOR	,894	,087	,009	,724	1,000
GREECE	,701	,041	,000	,621	,781
GUATEMALA	,853	,079	,009	,699	1,000
HAITI	1,000	,000	,064	1,000	1,000
INDIA	,937	,029	,000	,879	,994
INDONESIA	,663	,032	,000	,600	,726
IRAN	,866	,058	,030	,751	,980
ITALY	,643	,059	,016	,527	,758
JAPAN	,568	,041	,054	,488	,647
MEXICO	,681	,050	,002	,583	,778
NEW ZEALAND	,647	,062	,024	,525	,769
PAKISTAN	,904	,075	,009	,757	1,000
PANAMA	,964	,044	,004	,877	1,000
PERU	,734	,043	,000	,650	,818
PHILIPPINES	,592	,054	,088	,486	,699
PORTUGAL	,893	,086	,011	,725	1,000
SOLOMON ISLANDS	,361	,078	,090	,209	,514
SOUTH KOREA	,885	,125	,024	,640	1,000
TAIWAN	,649	,070	,044	,511	,786
TURKEY	,831	,033	,000	,766	,895
UK	1,000	,000	,053	1,000	1,000
UKRAINE	,917	,101	,053	,718	1,000
USA	,704	,036	,000	,634	,775
VANUATU	,314	,086	,034	,144	,483
VENEZUELA	,856	,076	,001	,707	1,000

Πίνακας 3.17 Ποσοστά ορθής ταξινόμησης ανά χώρα – Μέγεθος σεισμού

Discriminant Analysis			
Country	Youden's Maximum	Magnitude(c)	Ποσοστό Ορθής Ταξινόμησης (%)
ALBANIA	0,533	6,25	85,7
ALGERIA	0,806	5,65	74,6
AUSTRALIA	0,824	7,75	78,3
CHILE	0,16	6,85	55,6
CHINA	0,26	7,25	87,3
COSTA RICA	0,909	6,65	89,6
CROATIA	1	6,25	80,7
CUBA	0	8,9	60
ECUADOR	0,6	7,85	91,5
EGYPT	0,75	5,6	87,6
EL SALVADOR	0,667	6,95	94,2
GREECE	0,4	6,45	69,6
GUATEMALA	0,69	7,25	86,1
HAITI	1	6,35	60
INDIA	0,787	6,25	79,3
INDONESIA	0,284	6,45	55,2
IRAN	0,696	6,45	80,6
ITALY	0,29	6,55	71,5
JAPAN	0,285	6,35	76,6
MEXICO	0,367	6,85	70,6
NEW ZEALAND	0,24	6,85	55,7
PAKISTAN	0,693	7,45	93,8
PANAMA	0,909	6,75	84,2
PERU	0,379	6,45	58,3
PHILIPPINES	0,222	5,9	71,1
PORTUGAL	0,75	6,65	66,1
SOLOMON ISLANDS	0,093	7,85	42,1
SOUTH KOREA	0,75	6,4	88,9
TAIWAN	0,259	7,05	71,1
TURKEY	0,619	6,35	76,7
UK	1	5,4	73,3
UKRAINE	0,667	5,85	70
USA	0,374	6,25	59,7
VANUATU	0,095	5,95	50
VENEZUELA	0,721	6,4	79

Στον παραπάνω πίνακα φαίνεται πως για το Ελ Σαλβαδόρ μέγεθος σεισμού της τάξης των 6,95 ρίχτερ είναι ικανό να προκαλέσει τσουνάμι με ποσοστό ορθής ταξινόμησης 94,2%. Δηλαδή το συγκεκριμένο μέγεθος σεισμού μπορεί να χρησιμοποιηθεί για την πρόβλεψη φαινομένων τσουνάμι, ως κάτω άκρο, αρκετά ικανοποιητικά.

Χρησιμοποιώντας την ένταση του σεισμού ως συνεχή μεταβλητή, παρόλο που δεν είναι, τα αποτελέσματα φαίνονται παρακάτω.

Πίνακας 3.18 AUC ανά γεωγραφική περιοχή – Ένταση σεισμού

Area Under the Curve					
Test Result Variable(s):eq_intensity					
eq_region_code	Area	Std. Error ^a	AsymptoticSig. ^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
Caribbean	,669	,091	,086	,490	,847
Southern Europe	,402	,028	,000	,347	,456
Middle East	,652	,048	,002	,558	,745

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Πίνακας 3.19 Ποσοστά ορθής ταξινόμησης ανά γεωγραφική περιοχή – Ένταση σεισμού

Discriminant Analysis			
Region	Youden's Maximum	Mercalli(c)	Ποσοστό Ορθής Ταξινόμησης (%)
Caribbean	0,325	8,5	52
Middle East	0,264	8,5	76,5

Όσον αφορά τις ευρύτερες γεωγραφικές περιοχές τα αποτελέσματα δεν είναι καθόλου ικανοποιητικά όπως φαίνεται και στους παραπάνω πίνακες.

Πίνακας 3.20 AUC ανά χώρα – Ένταση σεισμού

Area Under the Curve					
Test Result Variable(s):eq_intensity					
country	Area	Std. Error ^a	AsymptoticSig. ^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
CROATIA	,988	,018	,099	,952	1,000
ITALY	,266	,042	,000	,185	,348
PAKISTAN	,917	,115	,077	,691	1,000
SYRIA	,809	,114	,060	,586	1,000

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Πίνακας 3.21 Ποσοστά ορθής ταξινόμησης ανά χώρα – Ένταση σεισμού

Discriminant Analysis			
Country	Youden's Maximum	Mercalli(c)	Ποσοστό Ορθής Ταξινόμησης (%)
CROATIA	0,976	10,5	77,2
PAKISTAN	0,75	6,5	91,7
SYRIA	0,456	9,5	76,5

Στην κατά χώρα ανάλυση παρατηρούμε ότι για το Πακιστάν ένταση σεισμού της τάξης του 7 στη τροποποιημένη κλίμακα Mercalli είναι ικανή να προκαλέσει τσουνάμι με ποσοστό ορθής ταξινόμησης 91,7% που είναι αρκετά ικανοποιητικό. Παρόλα αυτά στην ανάλυση με το μέγεθος του σεισμού για το Πακιστάν το ποσοστό ορθής ταξινόμησης ήταν 93,8% που κρίνεται προτιμότερο.

Επειδή στην πλειοψηφία τους τα ποσοστά ορθής ταξινόμησης στις παραπάνω χώρες και ευρύτερες γεωγραφικές περιοχές δεν ήταν ικανοποιητικά και δεδομένου ότι σε προηγούμενες παραγράφους υπήρχαν ενδείξεις ότι το εστιακό βάθος του σεισμού είναι μια σημαντική παράμετρος στην πρόβλεψη ενός φαινομένου τσουνάμι, πραγματοποιήθηκε αντίστοιχη ανάλυση για επιφανειακούς σεισμούς, δηλαδή σεισμούς με εστιακό βάθος μικρότερο των 60 χιλιομέτρων.

Πίνακας 3.22 AUC επιφανειακών σεισμών ανά γεωγραφική περιοχή – Μέγεθος σεισμού

Area Under the Curve (focal depth<60km)					
Test Result Variable(s):eq_primary					
eq_region_code	Area	Std. Error ^a	AsymptoticSig. ^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
Northern Africa	,901	,052	,001	,798	1,000
East Asia	,802	,023	,000	,757	,846
S. and SE. Asia and Indian Ocean	,921	,023	,000	,876	,965
Caribbean	,728	,085	,012	,562	,894
Central America	,915	,035	,000	,847	,983
Southern Europe	,767	,048	,000	,674	,860
Middle East	,845	,078	,001	,693	,998
North America and Hawaii	,784	,032	,000	,722	,847
South America	,797	,027	,000	,743	,850
Central and South Pacific	,663	,025	,000	,615	,712

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Πίνακας 3.23 Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά γεωγραφική περιοχή – Μέγεθος σεισμού

Discriminant Analysis			
Region	Youden's Maximum	Magnitude(c)	Ποσοστό Ορθής Ταξινόμησης (%)
Northern Africa	0,786	5,65	82,9
East Asia	0,571	6,25	76,6
S. and SE. Asia and Indian Ocean	0,722	6,45	79,4
Caribbean	0,482	6,75	68,1
Central America	0,833	6,75	86,1
Southern Europe	0,474	6,25	80,6
Middle East	0,577	6,55	82,7
North America and Hawaii	0,491	6,75	73,7
South America	0,511	6,85	72,6
Central and South Pacific	0,29	6,75	64,3

Όσον αφορά τις ευρύτερες γεωγραφικές περιοχές με έντονη σεισμική δραστηριότητα δεν παρατηρήθηκε κάποιο ικανοποιητικό ποσοστό ορθής ταξινόμησης στα μεγέθη των επιφανειακών σεισμών όπως φαίνεται στους παραπάνω πίνακες.

Πίνακας 3.24 AUC επιφανειακών σεισμών ανά χώρα – Μέγεθος σεισμού

Area Under the Curve (focal depth<60km)					
Test Result Variable(s):eq_primary					
country	Area	Std. Error^a	AsymptoticSig.^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
ALGERIA	,925	,058	,008	,811	1,000
AUSTRALIA	,979	,032	,010	,917	1,000
CHILE	,695	,069	,012	,561	,830
CHINA	,740	,167	,067	,413	1,000
COLOMBIA	1,000	,000	,005	1,000	1,000
COSTA RICA	,960	,048	,005	,865	1,000
ECUADOR	,865	,109	,090	,652	1,000
EGYPT	1,000	,000	,083	1,000	1,000
EL SALVADOR	1,000	,000	,053	1,000	1,000
GREECE	,759	,062	,000	,637	,882
GUATEMALA	,889	,079	,038	,735	1,000
INDIA	,960	,026	,000	,910	1,000
INDONESIA	,760	,037	,000	,689	,832
ITALY	,679	,100	,067	,482	,875
JAPAN	,653	,053	,001	,549	,758
MEXICO	,733	,061	,002	,614	,853
NEW ZEALAND	,740	,063	,001	,615	,864
PAKISTAN	,918	,069	,007	,783	1,000
PANAMA	,933	,074	,028	,789	1,000
PERU	,812	,045	,000	,724	,900
PHILIPPINES	,660	,061	,014	,540	,781
PORTUGAL	1,000	,000	,053	1,000	1,000
TAIWAN	,746	,085	,013	,580	,912
TURKEY	,900	,056	,003	,790	1,000
USA	,828	,036	,000	,757	,899
VENEZUELA	,943	,057	,001	,832	1,000

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Πίνακας 3.25 Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά χώρα – Μέγεθος σεισμού

Discriminant Analysis			
Country	Youden's Maximum	Magnitude(c)	Ποσοστό Ορθής Ταξινόμησης (%)
ALGERIA	0,85	5,65	87,5
AUSTRALIA	0,875	7,8	89,5
CHILE	0,321	6,85	70,2
CHINA	0,621	6,75	72,1
COLOMBIA	1	7,55	100
COSTA RICA	0,9	6,65	93,3
ECUADOR	0,731	6,85	75,9
EGYPT	1	6,15	100
EL SALVADOR	1	6,95	100
GREECE	0,463	6,25	70,5
GUATEMALA	0,867	7,25	88,9
INDIA	0,839	6,4	85,9
INDONESIA	0,425	6,45	67,7
ITALY	0,332	5,85	68,4
JAPAN	0,374	6,15	80,6
MEXICO	0,475	6,85	67,2
NEW ZEALAND	0,406	6,85	70
PAKISTAN	0,719	7,45	94,6
PANAMA	0,9	6,75	85,7
PERU	0,562	7,35	73,4
PHILIPPINES	0,3	6,75	65,1
PORTUGAL	1	6,55	100
TAIWAN	0,462	7,05	76,9
TURKEY	0,68	6,05	69,6
USA	0,57	6,25	76,8
VENEZUELA	0,933	6,4	91,3

Όπως φαίνεται στους παραπάνω πίνακες στην ανάλυση των επιφανειακών σεισμών, για κάθε χώρα ξεχωριστά, παρατηρήθηκε ότι για τις χώρες Κολομβία, Αίγυπτο, Ελ Σαλβαδόρ και Πορτογαλία τα μεγέθη των σεισμών που υποδείχθηκαν από τη μεγιστοποίηση της συνάρτησης του Youden ταξινομούν σωστά με 100% τα υπάρχοντα δεδομένα, πράγμα που σημαίνει ότι αυτά τα μεγέθη μπορούν να χρησιμοποιηθούν για την πρόβλεψη φαινομένων τσουνάμι με πολύ ικανοποιητικά αποτελέσματα.

Πίνακας 3.26 AUC επιφανειακών σεισμών ανά χώρα – Ένταση σεισμού

Area Under the Curve (focal depth<60km)					
Test Result Variable(s):eq_intensity					
country	Area	Std. Error ^a	AsymptoticSig. b	Asymptotic 95% ConfidenceInterval	
				LowerBound	UpperBound
CHILE	,743	,092	,040	,561	,924
CHINA	,780	,091	,064	,602	,958
MEXICO	,756	,119	,060	,523	,988
PHILIPPINES	,783	,066	,001	,655	,912

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Πίνακας 3.27 Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά χώρα – Ένταση σεισμού

Discriminant Analysis			
Country	Youden's Maximum	Mercalli(c)	Ποσοστό Ορθής Ταξινόμησης (%)
CHILE	0,481	6,5	41,6
CHINA	0,540	8,5	88,4
MEXICO	0,567	7,5	70,3
PHILIPPINES	0,4	6,5	71,3

Πίνακας 3.28 AUC επιφανειακών σεισμών ανά γεωγραφική περιοχή – Ένταση σεισμού

Area Under the Curve (focal depth<60km)					
Test Result Variable(s):eq_intensity					
eq_region_code	Area	Std. Error ^a	AsymptoticSig. b	Asymptotic 95% ConfidenceInterval	
				LowerBound	UpperBound
S. and SE. Asia and Indian Ocean	,694	,121	,098	,457	,931
Southern Europe	,610	,057	,061	,498	,721
Middle East	,819	,103	,017	,616	1,000
Central and South Pacific	,584	,049	,088	,489	,680

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Πίνακας 3.29 Ποσοστά ορθής ταξινόμησης επιφανειακών σεισμών ανά γεωγραφική περιοχή – Ένταση σεισμού

Discriminant Analysis			
Region	Youden's Maximum	Mercalli(c)	Ποσοστό Ορθής Ταξινόμησης (%)
S. and SE. Asia and Indian Ocean	0,541	9,5	89,4
Southern Europe	0,198	8,5	68,9
Middle East	0,613	9,5	89,2
Central and South Pacific	0,174	6,5	53,9

Στους τέσσερις παραπάνω πίνακες φαίνεται ότι η ένταση του σεισμού δεν είναι ικανή μέσω αυτής της στατιστικής μεθοδολογίας να αποδώσει ικανοποιητικά αποτελέσματα.

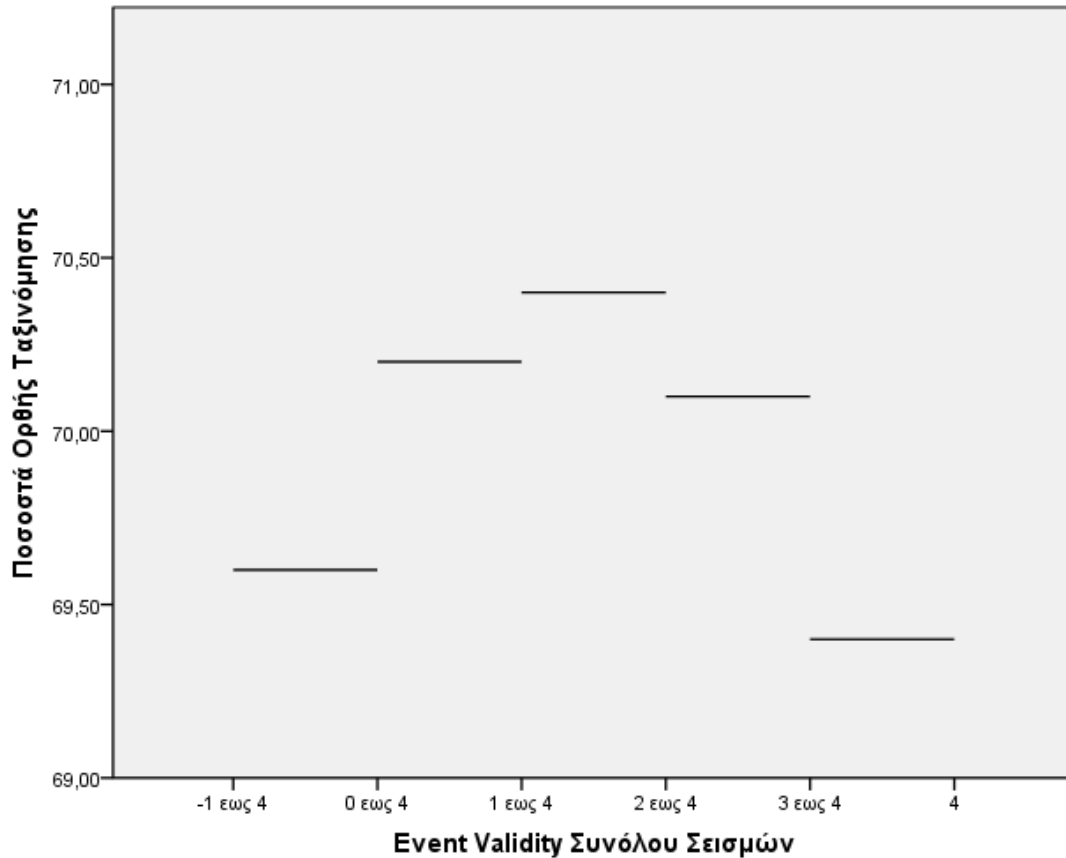
Πρέπει να ληφθεί υπόψιν ότι το σετ δεδομένων δεν είναι το καλύτερο δυνατό εφόσον οι καταγραφές των τσουνάμι είναι ως επί το πλείστον αμφίβολες. Παρακάτω παρατίθεται πίνακας με την αξιολόγηση 141 καταγραφών τσουνάμι με βάση τη μεταβλητή event_validity.

Πίνακας 3.30 Πίνακας συνάφειας καταγραφών τσουνάμι με την αξιολόγησή τους

event_validity * flag_tsunami Crosstabulation					
		flag_tsunami			
		Tsunami occurred		Total	
event_validity	erroneousentry	Count	11	11	
		% of Total	7,8%	7,8%	
	event that only caused a seiche or disturbance in an inland river	Count	1	1	
		% of Total	,7%	,7%	
	verydoubtfulsunami	Count	33	33	
		% of Total	23,4%	23,4%	
	questionabletsunami	Count	44	44	
		% of Total	31,2%	31,2%	
	probabletsunami	Count	16	16	
		% of Total	11,3%	11,3%	
	definitetsunami	Count	36	36	
		% of Total	25,5%	25,5%	
	Total	Count	141	141	
		% of Total	100,0%	100,0%	

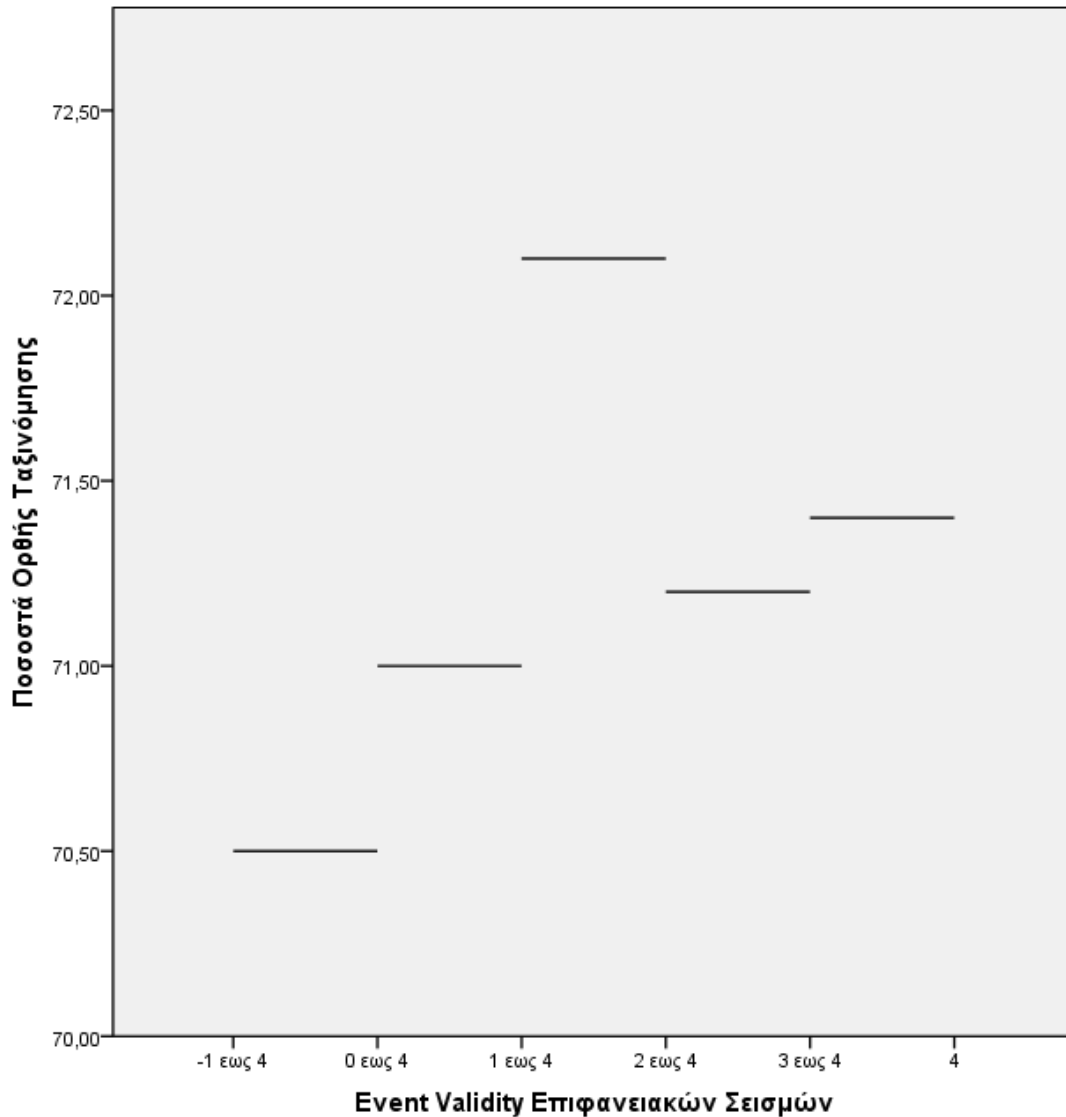
Στον παραπάνω πίνακα φαίνεται ότι στις 141 καταγραφές τσουνάμι μόνο οι 36 είναι σίγουρα σωστές, οι 16 είναι πιθανά τσουνάμι, οι 44 αμφίβολα τσουνάμι ενώ οι υπόλοιπες είτε είναι λανθασμένες καταγραφές είτε ως τσουνάμι καταγράφηκε μια απλή διαταραχή του νερού σε ποτάμι.

Εικόνα 3.3 Γράφημα αξιοπιστίας συνόλου σεισμών



Στο παραπάνω γράφημα φαίνεται η μεταβολή των ποσοστών ορθής ταξινόμησης στο σύνολο των σεισμών για την Ελλάδα αφαιρώντας κάθε φορά μία κατηγορία της μεταβλητής event validity. Το αξιοπερίεργο είναι ότι ενώ αναμενόταν το ποσοστό να αυξάνει συνεχώς υπάρχει αύξηση και μετά μείωση καταλήγοντας στο επίπεδο του 67,3% χρησιμοποιώντας μόνο τα σίγουρα τσουνάμι που είναι μικρότερο του αρχικά υπολογισμένου επιπέδου (<69,4%).

Εικόνα 3.4 Γράφημα αξιοπιστίας επιφανειακών σεισμών



Παρόμοια εικόνα παρουσιάζεται και χρησιμοποιώντας μόνο τους σεισμούς με εστιακό βάθος μικρότερο των 60km, όμως η χρήση μόνο των σίγουρων τσουνάμι δίνει ποσοστό ορθής ταξινόμησης 71,9% αυξημένο από το αρχικά υπολογισμένο επίπεδο κατά 1,4 μονάδες.

Η συγκεκριμένη ανωμαλία παρουσιάζεται εν μέρει λόγω των λανθασμένων πληροφοριών που υπάρχουν στο σετ δεδομένων όσων αφορά την Ελλάδα, αλλά σημαντικό πόλο παίζει και η διαφορετικότητα του Ιονίου από το Αιγαίο πέλαγος όσων αφορά τη μορφολογία και τα μεγέθη σεισμών που συγκεντρώνουν.

3.8 Χωρική Ανάλυση

Για τη διερεύνηση γεωγραφικών περιοχών ή χωρών με μεγαλύτερο κίνδυνο τσουνάμι από τις υπόλοιπες πραγματοποιήθηκε χωρική ανάλυση των δεδομένων στο στατιστικό πακέτο SatScan 9.4.1. Για τις χώρες χρησιμοποιήθηκαν το γεωγραφικό μήκος και πλάτος των πρωτευουσών τους ενώ για τις ευρύτερες γεωγραφικές περιοχές το γεωγραφικό μήκος και πλάτος του κέντρου κάθε περιοχής.

Οι αναλύσεις που ακολουθούν είναι καθαρά χωρικές οι οποίες ανιχνεύουν συστάδες με υψηλά ποσοστά εμφάνισης τσουνάμι και υποθέτουμε ότι οι εμφανίσεις τσουνάμι ακολουθούν την διακριτή κατανομή Poisson με σταθερό ρυθμό.

SaTScan v9.1.1

Program run on: Tue Mar 24 11:22:49 2015

Purely Spatial analysis

scanning for clusters with high rates

using the Discrete Poisson model.

SUMMARY OF DATA

Study period.....: 1900/1/1 to 2020/12/31

Number of locations.....: 149

Total population.....: 6088

Total number of cases.....: 1985
Annual cases / 100000.....: 269.5

MOST LIKELY CLUSTER

1.Location IDs included.: Wellington, Noumea, Canberra,
Nuku'alofa, Suva, Port_Vila, Avarua,
Apia, Antarctica, Honiara, Papeete,
Port_Moresby, Tarawa, Palikir,
Melekeok, Hagatna, Pacific_Ocean,
Usa_terittory, Jakarta, Indian_Ocean,
Manila, Kuala_Lumpur,
King_Edward_Point, Taipei, Tokyo,
Santiago
Coordinates / radius..: (41.300000 S, 174.783333 E) / 9342.00 km
Population.....: 1664
Number of cases.....: 973
Expected cases.....: 542.55
Annual cases / 100000.: 483.2
Observed / expected...: 1.79
Relative risk.....: 2.56
Log likelihood ratio..: 209.666361
P-value.....: < 0.000000000000000010

SECONDARY CLUSTERS

2.Location IDs included.: Washington_D_C, Ottawa, Havana,
Port_au_Prince, Kingston

Coordinates / radius...: (38.883333 N, 77.000000 W) / 2320.74 km
Population.....: 322
Number of cases.....: 163
Expected cases.....: 104.99
Annual cases / 100000.: 418.4
Observed / expected...: 1.55
Relative risk.....: 1.60
Log likelihood ratio...: 14.596479
P-value.....: 0.000038

3.Location IDs included.: Moscow, Kyiv, Warsaw, Stockholm

Coordinates / radius...: (55.750000 N, 37.600000 E) / 1226.74 km
Population.....: 162
Number of cases.....: 92
Expected cases.....: 52.82
Annual cases / 100000.: 469.3
Observed / expected...: 1.74
Relative risk.....: 1.78
Log likelihood ratio...: 12.270398
P-value.....: 0.00034

4.Location IDs included.: Athens

Coordinates / radius...: (37.983333 N, 23.733333 E) / 0 km
Population.....: 306

Number of cases.....: 139
Expected cases.....: 99.77
Annual cases / 100000.: 375.4
Observed / expected...: 1.39
Relative risk.....: 1.42
Log likelihood ratio..: 7.273625
P-value.....: 0.039

PARAMETER SETTINGS

Input

Case File : D:\Capitals.CAS
Population File : D:\Capitals.pop
Coordinates File : D:\Capitals.GEO
Time Precision : Year
Start Time : 1900/1/1
End Time : 2020/12/31
Coordinates : Latitude/Longitude

Analysis

Type of Analysis : Purely Spatial
Probability Model : Discrete Poisson
Scan for Areas with : High Rates

Output

Results File : D:\fsdfds

Data Checking

Temporal Data Check : Check to ensure that all cases and controls are within the specified temporal study period.

Geographical Data Check : Check to ensure that all observations (cases, controls and populations) are within the specified geographical area.

Spatial Neighbors

Use Non-Euclidian Neighbors file : No

Use Meta Locations File : No

Multiple Coordinates Type : Allow only set of coordinates per location ID.

Spatial Window

Maximum Spatial Cluster Size : 50 percent of population at risk

Window Shape : Circular

Isotonic Scan : No

Space And Time Adjustments

Adjust for known relative risks : No

Inference

P-Value Reporting : Default Combination
Adjusting for More Likely Clusters : No
Number of Replications : 999

Clusters Reported

Criteria for Reporting Secondary Clusters : No Geographical Overlap

Additional Output

Report Critical Values : No
Report Monte Carlo Rank : No

Run Options

Processor Usage : All Available Processors
Logging Analysis : Yes
Suppress Warnings : No

Program completed : Tue Mar 24 11:22:49 2015

Total Running Time : 0 seconds

Processor Usage : 8 processors

Όσον αφορά τις πρωτεύουσες υπάρχουν 149 πόλεις, συνολικές περιπτώσεις σεισμών 6088 και συνολικές εμφανίσεις τσουνάμι 1985.

Πίνακας 3.31 Συστάδες ανά χώρα

Συστάδα	Capitals included	Population	Cases					
			Observed	Expected	Ratio	Relative Risk	Log-likelihood ratio	p-value
1	Wellington, Noumea, Canberra, Nuku'alofa, Suva, Port_Vila, Avarua, Apia, Antarctica, Honiara, Papeete, Port_Moresby, Tarawa, Palikir, Melekeok, Hagatna, Pacific_Ocean, Usa_territory, Jakarta, Indian_Ocean, Manila, Kuala_Lumpur, King_Edward_Point, Taipei, Tokyo, Santiago	1664	973	542,55	1,79	2,56	209,67	<1‰
2	Washington_D_C, Ottawa, Havana, Port_au_Prince, Kingston	322	163	104,99	1,55	1,6	14,596	<1‰
3	Moscow, Kyiv, Warsaw, Stockholm	162	92	52,82	1,74	1,78	12,27	<1‰
4	Athens	306	139	99,77	1,39	1,42	7,27	0,039

Η πιο πιθανή συστάδα περιλαμβάνει τις πρωτεύουσες Wellington, Noumea, Canberra, Nuku'alofa, Suva, Port Vila, Avarua, Apia, Antarctica, Honiara, Papeete, Port Moresby, Tarawa, Palikir, Melekeok, Hagatna, Pacific Ocean, Usa territory, Jakarta, Indian Ocean, Manila, Kuala Lumpur, King Edward Point, Taipei, Tokyo και Santiago. Ο συνολικός αριθμός σεισμών είναι 1664 με συνολικές εμφανίσεις τσουνάμι 973. Ο αναμενόμενος αριθμός εμφανίσεων τσουνάμι για τη συγκεκριμένη συστάδα, αν οι εμφανίσεις τσουνάμι παγκοσμίως ήταν τυχαία διασκορπισμένες στο χάρτη και ακολουθούσαν μία κοινή κατανομή Poisson, είναι 542,55. Ουσιαστικά το πρόγραμμα κατασκευάζει ένα πίνακα συνάφειας για να υπολογίσει τις αναμενόμενες εμφανίσεις τσουνάμι. Ο πίνακας είναι:

	Συστάδα	Υπολοιπος Κόσμος	Σύνολο
Εμφάνιση τσουνάμι	973	1012	1985
Μη εμφάνιση τσουνάμι	691	3412	4103
Σύνολο	1664	4424	6088

Αναμενόμενη Τιμή πρώτου κελιού

$$E = \frac{X_{1.} \cdot X_{.1}}{X_{..}} = \frac{1985 \cdot 1664}{6088} = 542,5493$$

Τα ratio είναι ο λόγος των συνολικών εμφανίσεων προς τις αναμενόμενες εμφανίσεις, που είναι ίσος με 1,79, το οποίο σημαίνει ότι σε αυτή τη συστάδα υπάρχει 79% μεγαλύτερος κίνδυνος εμφάνισης τσουνάμι παγκοσμίως. Ο σχετικός κίνδυνος υπολογίστηκε 2,56 που σημαίνει ότι ο κίνδυνος να εμφανιστεί τσουνάμι μετά από σεισμό είναι περίπου 2,5 φορές τον “κίνδυνο” να μην εμφανιστεί τσουνάμι μετά από σεισμό. Η αύξηση του κινδύνου κρίνεται στατιστικά σημαντική εφόσον η p-value του ελέγχου είναι μικρότερη του 1%.

Στις δευτερεύουσες συστάδες παρατηρούμε ότι η Αθήνα είναι μια συστάδα μόνη της με συνολικούς σεισμούς 306, εμφανίσεις τσουνάμι 139 και αναμενόμενες εμφανίσεις 99,77. Στην Αθήνα υπάρχει 39% μεγαλύτερος κίνδυνος εμφάνισης τσουνάμι σε σχέση με τον υπόλοιπο κόσμο. Ο σχετικός κίνδυνος υπολογίστηκε 1,42 που σημαίνει ότι ο κίνδυνος να εμφανιστεί τσουνάμι μετά από σεισμό είναι περίπου 1,5 φορά τον “κίνδυνο” να μην εμφανιστεί τσουνάμι μετά από σεισμό. Η αύξηση του κινδύνου κρίνεται στατιστικά σημαντική εφόσον η p-value του ελέγχου είναι 0,039.

SaTScan v9.4.1

Program run on: Tue Mar 24 20:04:44 2015

Purely Spatial analysis
scanning for clusters with high rates
using the Discrete Poisson model.

SUMMARY OF DATA

Study period.....: 1900/1/1 to 2020/12/31
Number of locations.....: 18
Total population.....: 5839
Total number of cases.....: 1736

Annual cases / 100000.....: 245.7

CLUSTERS DETECTED

1.Location IDs included.: Bering_Sea, North_America_and_Hawaii,
Central_and_South_Pacific, Central_America, Caribbean, South_America,
Atlantic_Ocean

Overlap with clusters.: 2, 4, 5
Coordinates / radius...: (56.9073,-178.139) / 159.94
Gini Cluster.....: No
Population.....: 2233
Number of cases.....: 969
Expected cases.....: 663.90
Annual cases / 100000.: 358.6
Observed / expected...: 1.46
Relative risk.....: 2.04
Log likelihood ratio...: 109.555065
P-value.....: < 0.000000000000000001

2.Location IDs included.: Central_and_South_Pacific

Overlap with clusters.: 1
Coordinates / radius...: (-6.51854,-145.491) / 0
Gini Cluster.....: Yes
Population.....: 892
Number of cases.....: 462
Expected cases.....: 265.20
Annual cases / 100000.: 428.0
Observed / expected...: 1.74
Relative risk.....: 2.01
Log likelihood ratio...: 73.442253
P-value.....: < 0.000000000000000001

3.Location IDs included.: Kamchatka_and_Kuril_Islands

Overlap with clusters.: No Overlap

Coordinates / radius...: (47.596,153.505) / 0
Gini Cluster.....: Yes
Population.....: 118
Number of cases.....: 78
Expected cases.....: 35.08
Annual cases / 100000.: 546.3
Observed / expected...: 2.22
Relative risk.....: 2.28
Log likelihood ratio...: 19.950783
P-value.....: 0.0000000036

4.Location IDs included.: Atlantic_Ocean, South_America, Caribbean

Overlap with clusters.: 1
Coordinates / radius...: (-7.20086,-31.608) / 50.88
Gini Cluster.....: Yes
Population.....: 730
Number of cases.....: 295
Expected cases.....: 217.04
Annual cases / 100000.: 334.0
Observed / expected...: 1.36
Relative risk.....: 1.43
Log likelihood ratio...: 14.610568
P-value.....: 0.0000011

5.Location IDs included.: Bering_Sea, North_America_and_Hawaii

Overlap with clusters.: 1
Coordinates / radius...: (56.9073,-178.139) / 62.48
Gini Cluster.....: Yes
Population.....: 460
Number of cases.....: 183
Expected cases.....: 136.76
Annual cases / 100000.: 328.8
Observed / expected...: 1.34
Relative risk.....: 1.38

Log likelihood ratio...: 7.734191
P-value.....: 0.0016

PARAMETER SETTINGS

Input

Case File : C:\Users\panos\Desktop\Region.cas
Population File : C:\Users\panos\Desktop\Region.pop
Time Precision : Year
Start Time : 1900/1/1
End Time : 2020/12/31
Coordinates File : C:\Users\panos\Desktop\Region.geo
Coordinates : Cartesian

Analysis

Type of Analysis : Purely Spatial
Probability Model : Discrete Poisson
Scan for Areas with : High Rates

Output

Main Results File : C:\Users\panos\Desktop\reg.txt

Data Checking

Temporal Data Check : Check to ensure that all cases and controls are within the specified temporal study period.

Geographical Data Check : Check to ensure that all observations (cases, controls and populations) are within the specified geographical area.

Spatial Neighbors

Use Non-Euclidian Neighbors file : No
Use Meta Locations File : No
Multiple Coordinates Type : Allow only one set of
coordinates per location ID.

Spatial Window

Maximum Spatial Cluster Size : 50 percent of population at risk
Window Shape : Circular
Isotonic Scan : No

Space And Time Adjustments

Adjust for known relative risks : No

Inference

P-Value Reporting : Default Combination
Number of Replications : 999
Adjusting for More Likely Clusters : No

Border Analysis

Report Oliveira's F : No

Spatial Output

Report Hierarchical Clusters : Yes
Criteria for Reporting Secondary Clusters : No Geographical Overlap
Report Gini Optimized Cluster Collection : Yes
Gini Index Based Collection Reporting : Optimal Only
Report Gini Index Cluster Coefficients : No
Spatial Cluster Maxima : 1, 2, 3, 4, 5, 6, 8,
10, 12, 15, 20, 25, 30, 40, 50

Restrict Reporting to Smaller Clusters : No

Other Output

Report Critical Values : No

Report Monte Carlo Rank : No

Run Options

Processor Usage : All Available Processors

Suppress Warnings : No

Logging Analysis : Yes

Program completed : Tue Mar 24 20:04:44 2015

Total Running Time : 0 seconds

Processor Usage : 4 processors

Όσον αφορά τις ευρύτερες γεωγραφικές περιοχές υπάρχουν 18 με συνολικές περιπτώσεις σεισμών 5839 και συνολικές εμφανίσεις τσουνάμι 1736.

Πίνακας 3.32 Συστάδες ανά γεωγραφική περιοχή

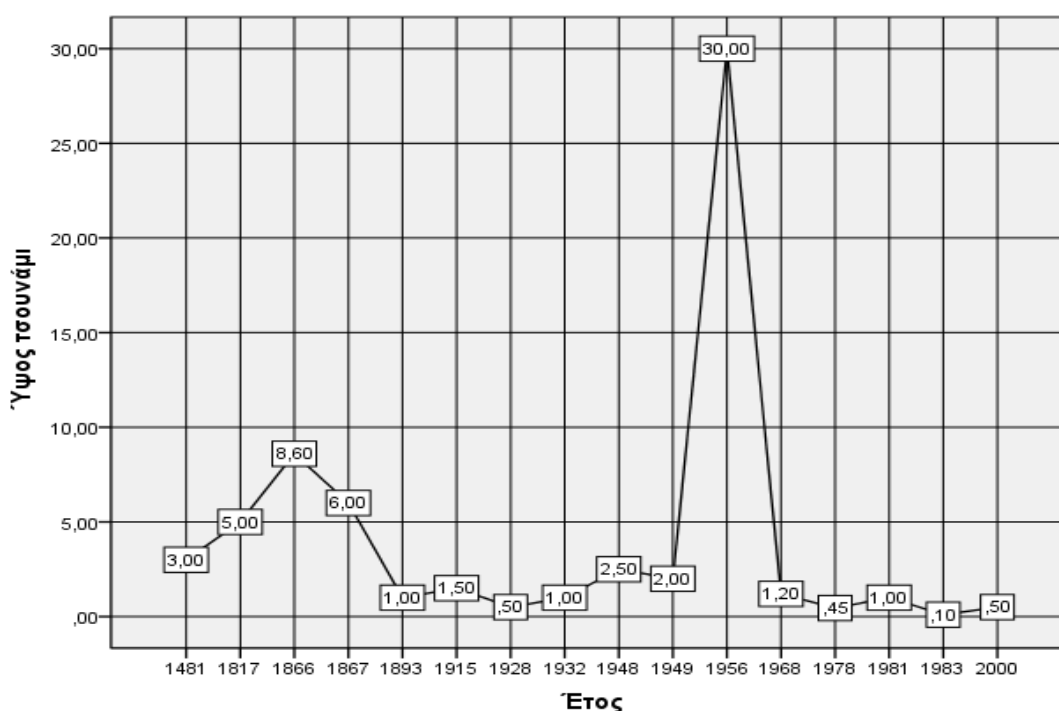
Συστάδα	Regions included	Population	Cases					
			Observed	Expected	Ratio	Relative Risk	Log-likelihood ratio	p-value
1	Bering Sea, North America and Hawaii, Central and South Pacific, Central America, Caribbean, South America, Atlantic Ocean	2233	969	663.9	1.46	2.04	109.56	<1‰
2	Central and South Pacific	892	462	265.2	1.74	2.01	73.44	<1‰
3	Kamchatka and Kuril Islands	118	78	35.08	2.22	2.28	19.951	<1‰
4	Atlantic Ocean, South America, Caribbean	730	295	217.04	1.36	1.43	14.612	<1‰
5	Bering Sea, North America and Hawaii	460	183	136.76	1.34	1.38	7.734	<1‰

Δεν υπάρχει πιο πιθανή συστάδα οι μόνες που ξεχωρίζουν είναι αυτές που παρατίθενται στον παραπάνω πίνακα. Η πρώτη συστάδα περιλαμβάνει τις περιοχές Bering Sea, North America and Hawaii, Central and South Pacific, Central America, Caribbean, South America και Atlantic Ocean. Ο συνολικός αριθμός σεισμών είναι 2233 με συνολικές εμφανίσεις τσουνάμι 969. Ο αναμενόμενος αριθμός εμφανίσεων τσουνάμι με βάση την κατανομή Poisson είναι 663,9. Τα ratio είναι ίσο με 1,46, το οποίο σημαίνει ότι σε αυτή τη συστάδα υπάρχει 46% μεγαλύτερος κίνδυνος εμφάνισης τσουνάμι παγκοσμίως. Ο σχετικός κίνδυνος υπολογίστηκε 2,04 που σημαίνει ότι ο κίνδυνος να εμφανιστεί τσουνάμι μετά από σεισμό είναι περίπου 2 φορές τον “κίνδυνο” να μην εμφανιστεί τσουνάμι μετά από σεισμό. Η αύξηση του κινδύνου κρίνεται στατιστικά σημαντική εφόσον η p-value του ελέγχου είναι μικρότερη του 1%.

3.9 Περιοδικότητα

Ο δρ. Γεράσιμος Παπαδόπουλος, διευθυντής ερευνών του Γεωδυναμικού Ινστιτούτου, σε δήλωσή του στην εφημερίδα “Εθνος” στις 21/3/2015 βασισμένος στην περιοδικότητα του φαινομένου τσουνάμι, υποστήριξε ότι υπάρχει μεγάλη πιθανότητα εμφάνισης ενός τέτοιου φαινομένου μετά από επιφανειακό σεισμό μεγαλύτερο των 6 Richter.

Εικόνα 3.5 Γράφημα περιοδικότητας συνόλου τσουνάμι



Όπως παρατηρείται στο παραπάνω γράφημα δεν είναι ευδιάκριτη κάποια περιοδικότητα στο φαινόμενο. Επίσης για άλλη μια φορά αποδεικνύεται η αναξιοπιστία των δεδομένων εφόσον υπάρχουν καταγραφές τσουνάμι με ύψος κύματος ακόμα και 10 εκατοστών.

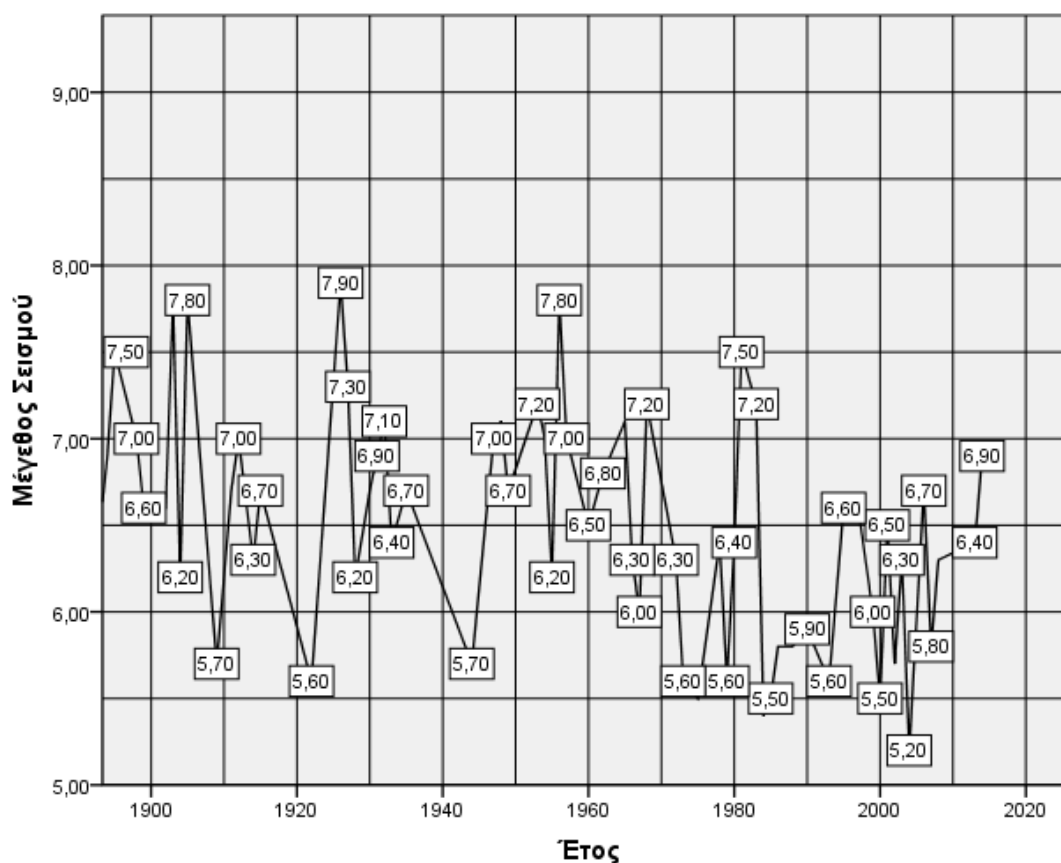
Εικόνα 3.6 Γράφημα περιοδικότητας επιβεβαιωμένων τσουνάμι



Χρησιμοποιώντας μόνο τις επιβεβαιωμένες καταγραφές του σετ δεδομένων δεν υπάρχει κάποια καλύτερη εικόνα για την περιοδικότητα του φαινομένου.

Συμπερασματικά, αν υποθέσουμε ότι οι δηλώσεις του κ. Παπαδόπουλου είναι ορθές, γίνεται προφανές ότι τα δεδομένα που υπάρχουν δημοσιευμένα σε ιστοσελίδες είναι ελλιπέστατα και ανακριβή.

Εικόνα 3.7 Γράφημα περιοδικότητας σεισμικής δραστηριότητας



Σε αντίθεση με τα δεδομένα για το φαινόμενο τσουνάμι, τα δεδομένων των σεισμών που υπάρχουν στο σύνολο δεδομένων φαίνεται να παρουσιάζουν περιοδικότητα (Εικόνα 3.7). Περίπου κάθε 20 χρόνια καταγράφεται ένας σεισμός άνω των 7 Richter στον ελλαδικό χώρο με εξαίρεση τις δύο τελευταίες εικοσαετίες που τα μέγιστα κυμαίνονται από 6,6 R με 6,9 R.

Κεφάλαιο 4.

Εμπόδια εκπόνησης και συμπεράσματα

Σε αυτό το κεφάλαιο συνοψίζονται τα αποτελέσματα όλης της στατιστικής ανάλυσης που εφαρμόστηκε στο σετ δεδομένων καθώς επίσης γίνεται μια εκτενής αναφορά στα εμπόδια που τέθηκαν από τους κρατικούς οργανισμούς για τη συλλογή των δεδομένων.

4.1 Εμπόδια εκπόνησης

Σε αντίθεση με όλες τις υπόλοιπες χώρες η εύρεση αξιόπιστου σετ δεδομένων για σεισμολογικά μεγέθη είναι πάρα πολύ δύσκολη. Η επαφή με το Γεωδυναμικό Ινστιτούτο και το Αστεροσκοπείο Αθηνών μπορεί μόνο να χαρακτηριστεί επιεικώς ως άκαρπη, όσον αφορά την αίτησή μας για πρόσβαση σε ένα πλήρες και αξιόπιστο σετ δεδομένων ώστε να γίνει εφικτή η ανάλυση του φαινομένου μόνο στον ελλαδικό χώρο χωρισμένη σε ρήγματα αλλά και σε πελάγη συνεκτιμώντας τις ιδιαιτερότητες κάθε τόπου. Η απάντηση που πήραμε στο αίτημα μας από υψηλά ιστάμενο εκπρόσωπο του Γεωδυναμικού Ινστιτούτου ήταν αρνητική.

Μετά από τέτοια αντιμετώπιση από τους ειδικούς στον τομέα της σεισμολογίας αναγκαστήκαμε να δημιουργήσουμε ένα σετ δεδομένων κατάλληλο για την στατιστική ανάλυση που είχαμε σχεδιάσει. Η κατασκευή αυτού του σετ δεδομένων περιλάμβανε στοιχεία από μία παγκόσμια ιστοσελίδα, που αναφέρεται παραπάνω, με τους σημαντικότερους σεισμούς που έχουν καταγραφεί καθώς και τα σημαντικότερα τσουνάμι που έχουν προκληθεί από σεισμό. Η αντιπαραβολή των 6000 περίπου παρατηρήσεων και η αντιστοίχιση των μεταβλητών που αφορούσαν τα τσουνάμι με τα γεγονότα και τις μεταβλητές των σεισμών ήταν μία επίπονη και χρονοβόρα διαδικασία που διήρκησε περίπου 10 ημέρες. Ακόμα και σε αυτό το σετ δεδομένων υπήρξε πρόβλημα στη χώρα γένεσης του σεισμού καθώς σεισμοί που το γεωγραφικό μήκος και πλάτος τους ανήκαν στα σύνορα της Ελλάδας είχαν καταχωρηθεί είτε στην Τουρκία είτε στην Ιταλία είτε στην Αίγυπτο. Το φαινόμενο αυτό παρατηρήθηκε και σε πολλές άλλες χώρες με αποτέλεσμα να πρέπει να διερευνηθούν τα γεωγραφικά μήκη και πλάτη όλων των παρατηρήσεων ως προς την ορθότητα της χώρας στην οποία έχουν καταχωρηθεί.

4.2 Συμπεράσματα-Προτάσεις

Όπως διαπιστώθηκε στο κεφάλαιο 3 το ύψος ενός κύματος τσουνάμι μπορεί να προβλεφθεί από ένα απλό γραμμικό στατιστικό μοντέλο γνωρίζοντας μόνο το μέγεθος του σεισμού και το εστιακό βάθος. Γνωρίζοντας τις ίδιες μεταβλητές θα μπορούσε να εκτιμηθεί η πιθανότητα εμφάνισης τσουνάμι με βάση ένα μοντέλο λογιστικής παλινδρόμησης. Ως εκ τούτου σε κάθε χώρα ξεχωριστά θα μπορούσε να γίνει η ανάλυση των δεδομένων που καταγράφονται από τους αντίστοιχους κρατικούς φορείς και έτσι να εφαρμοστεί ένα σύστημα πρόβλεψης γένεσης τσουνάμι και προειδοποίησης των περιοχών που θα πληγούν για αποφυγή των καταστροφικών συνεπειών του, χωρίς να χρειάζεται να δαπανηθούν ποσά στη κατασκευή σταθμών πολλά μίλια μακριά από την ακτή που για να βγάλουν προειδοποίηση χρειάζεται πρώτα να το δουν και να το καταγράψουν.

Επίσης για κάθε χώρα ξεχωριστά βρέθηκε ένα κάτω άκρο στο μέγεθος του σεισμού ικανό να προκαλέσει τσουνάμι είτε σε επιφανειακούς σεισμούς είτε γενικότερα και επομένως θα μπορούσε να υπάρχει προειδοποίηση για τσουνάμι πιο γρήγορα με αποτέλεσμα να μειωθεί στο ελάχιστο η θνησιμότητα του φαινομένου.

Με βάση τις τρεις παραπάνω αναλύσεις προτείνεται κάθε χώρα να επεξεργαστεί τα δεδομένα που καταγράφει για να εκτιμήσει ακριβέστερα είτε τα μοντέλα πρόβλεψης είτε το κάτω άκρο του μεγέθους του σεισμού που είναι ικανό να προκαλέσει τσουνάμι.

Τέλος με βάση τη χωρική ανάλυση που διενεργήθηκε, χώρες όπως η Ελλάδα, η Ιαπωνία, η Αμερική καθώς και άλλες, που εκτιμήθηκε ότι έχουν μεγαλύτερο κίνδυνο εμφάνισης τσουνάμι από τον υπόλοιπο κόσμο, προτείνεται να εφαρμόσουν όλη την παραπάνω ανάλυση καθώς και να εκτιμήσουν ένα μοντέλο διατακτικής παλινδρόμησης με εξαρτημένη μεταβλητή την ένταση του τσουνάμι και ανεξάρτητες το εστιακό βάθος και μέγεθος του σεισμού. Στο σετ δεδομένων που αναλύθηκε παραπάνω μια τέτοια ανάλυση ήταν αδύνατη για κάθε χώρα ξεχωριστά ή ακόμα και για κάθε ευρύτερη γεωγραφική περιοχή καθώς δεν υπήρχαν αρκετά δεδομένα. Πιθανόν αν υπήρχαν επαρκή δεδομένα για ανάλυση κατά χώρα να μπορούσαν να εκτιμηθούν ακριβέστερα τέτοια μοντέλα και έτσι όταν η πιθανότητα εμφάνισης τσουνάμι εκτιμάται υψηλή να μπορεί να γίνει πρόβλεψη και για τη σφοδρότητα με την οποία το κύμα θα χτυπήσει την ακτή.

ΠΑΡΑΡΤΗΜΑ Α

Variables Entered/Removed^{b, c}

Model	Variables Entered	Variables Removed	Method
1	focal_depth, eq_primary	.	Enter

a. All requested variables entered.

b. eq_region_code = East Asia

c. Dependent Variable: maximum_water_height

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,445 ^a	,198	,188	4,48329

a. Predictors: (Constant), focal_depth, eq_primary

b. eq_region_code = East Asia

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	750,858	2	375,429	18,678	,000 ^a
	Residual	3035,083	151	20,100		
	Total	3785,941	153			

a. Predictors: (Constant), focal_depth, eq_primary

b. Dependent Variable: maximum_water_height

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-21,269	3,834		-5,547	,000	-28,844	-13,693
	eq_primary	3,371	,552	,451	6,110	,000	2,281	4,461
	focal_depth	-,023	,021	-,081	-1,094	,276	-,065	,019

a. Dependent Variable: maximum_water_height

Στους παραπάνω πίνακες παρουσιάζονται τα αποτελέσματα της ανάλυσης παλινδρόμησης για την Ανατολική Ασία. Πιο αναλυτικά, ο συντελεστής προσδιορισμού είναι ίσος με 0,198 που σημαίνει ότι το 19,8% της μεταβλητότητας του μέγιστου ύψους κύματος ερμηνεύεται από το μέγεθος του σεισμού και το εστιακό βάθος. Αν και ο συντελεστής είναι αρκετά μικρός το μοντέλο κρίνεται στατιστικά σημαντικό εφόσον η p-value του F ελέγχου είναι μικρότερη του 1%. Από τον πίνακα των συντελεστών παρατηρούμε ότι ο σταθερός όρος και το μέγεθος του σεισμού

κρίνονται στατιστικά σημαντικά σε αντίθεση με το εστιακό βάθος του οποίου η p-value είναι 0,0276.

Διερευνώντας τις προϋποθέσεις της γραμμικής παλινδρόμησης παρατηρούμε ότι τα σφάλματα δεν ακολουθούν κανονική κατανομή.

One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
N		154
Normal Parameters ^{a, b}	Mean	,0000000
	Std. Deviation	4,45389117
Most Extreme Differences	Absolute	,276
	Positive	,276
	Negative	-,188
Kolmogorov-Smirnov Z		3,421
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Normal.
b. Calculated from data.

Ο έλεγχος κανονικότητας που χρησιμοποιήθηκε είναι ο Kolmogorov-Smirnov, ο οποίος είναι γνωστός για την τάση του να αποδέχεται πολύ εύκολα την κανονικότητα. Στο συγκεκριμένο παράδειγμα εφόσον ακόμα και αυτός ο έλεγχος απορρίπτει την κανονικότητα μπορούμε με βεβαιότητα να υποθέσουμε ότι δεν υπάρχει κανονικότητα.

Bootstrap Specifications

Sampling Method	Simple
Number of Samples	1000
Confidence Interval Level	95,0%
Confidence Interval Type	Percentile

Bootstrap for Coefficients

Model	B	Bootstrap ^a				
		Bias	Std. Error	Sig. (2-tailed)	95% Confidence Interval	
					Lower	Upper
1 (Constant)	-21,269	-,181	8,039	,055	-37,843	-7,797
eq_primary	3,371	,022	1,210	,052	1,282	5,819
focal_depth	-,023	,000	,014	,146	-,056	,000

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Η μέθοδος Bootstrap δημιούργησε 1000 δείγματα με επανατοποθέτηση από το αρχικό δείγμα των 154 παρατηρήσεων για την εκτίμηση των p-values των συντελεστών της παλινδρόμησης. Όπως παρατηρούμε ο σταθερός όρος και το μέγεθος του σεισμού

παραμένουν στατιστικά σημαντικά έστω και οριακά ενώ το εστιακό βάθος στατιστικά ασήμαντο όπως και πριν.

Χρησιμοποιώντας όλο το δείγμα των παρατηρήσεων εφαρμόσαμε μοντέλο λογιστικής παλινδρόμησης.

Model Information

Dependent Variable	flag_tsunami ^a
Probability Distribution	Binomial
Link Function	Logit

a. The procedure models Tsunami occurred as the response, treating Tsunami not occurred as the reference category.

Case Processing Summary

	N	Percent
Included	102	52,0%
Excluded	94	48,0%
Total	196	100,0%

Categorical Variable Information

			N	Percent
Dependent Variable	flag_tsunami	Tsunami not occurred	31	30,4%
		Tsunami occurred	71	69,6%
		Total	102	100,0%

Continuous Variable Information

		N	Minimum	Maximum	Mean	Std. Deviation
Covariate	eq_primary	102	5,00	9,50	7,1441	,77262
	focal_depth	102	4	180	45,66	29,666

Στους παραπάνω πίνακες παρατηρούμε κάποια περιγραφικά στοιχεία για τις συνεχείς μεταβλητές καθώς και το πλήθος των περιπτώσεων που υπάρχουν καταγραφές και για τις δύο ανεξάρτητες μεταβλητές (102).

Goodness of Fit^b

	Value	df	Value/df
Deviance	103,795	86	1,207
Scaled Deviance	103,795	86	
Pearson Chi-Square	98,034	86	1,140
Scaled Pearson Chi-Square	98,034	86	
Log Likelihood ^a	-52,590		
Akaike's Information Criterion (AIC)	111,181		
Finite Sample Corrected AIC (AICC)	111,426		
Bayesian Information Criterion (BIC)	119,056		
Consistent AIC (CAIC)	122,056		

Dependent Variable: flag_tsunami
 Model: (Intercept), eq_primary, focal_depth

- a. The full log likelihood function is displayed and used in computing information criteria.
- b. Information criteria are in small-is-better form.

Omnibus Test^a

Likelihood Ratio Chi-Square	df	Sig.
18,719	2	,000

Dependent Variable:
 flag_tsunami
 Model: (Intercept), eq_primary,
 focal_depth

- a. Compares the fitted model against the intercept-only model.

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	4,757	1	,029
eq_primary	8,597	1	,003
focal_depth	9,590	1	,002

Dependent Variable: flag_tsunami
 Model: (Intercept), eq_primary, focal_depth

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-5,242	2,4034	-9,953	-,531	4,757	1	,029	,005	4,761E-5	,588
eq_primary	1,040	,3548	,345	1,736	8,597	1	,003	2,830	1,412	5,674
focal_depth (Scale)	-,026 1 ^a	,0085	-,043	-,010	9,590	1	,002	,974	,958	,990

Dependent Variable: flag_tsunami
 Model: (Intercept), eq_primary, focal_depth

- a. Fixed at the displayed value.

Το μοντέλο κρίνεται στατιστικά σημαντικό εφόσον διαφέρει από το μοντέλο μόνο με το σταθερό όρο σύμφωνα με την p-value του Omnibus test. Στον έλεγχο των συντελεστών της λογιστικής παλινδρόμησης παρατηρούμε ότι και οι δύο ανεξάρτητες μεταβλητές καθώς και ο σταθερός όρος κρίνονται στατιστικά σημαντικά με p-values < 0,029.

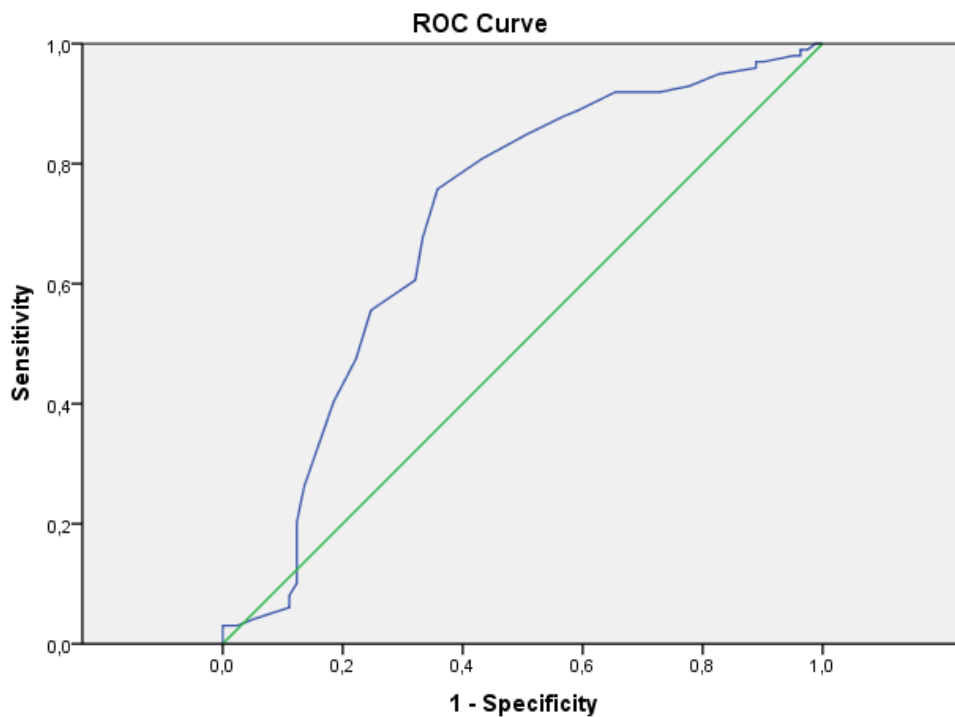
Χρησιμοποιώντας καμπύλες Roc για τα δεδομένα της Ελλάδας παρατηρήσαμε τα παρακάτω.

Case Processing Summary

flag_tsunami	Valid N (listwise)
Positive ^a	99
Negative	81
Missing	126

Larger values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is Tsunami occurred.



Diagonal segments are produced by ties.

Area Under the Curve

Test Result Variable(s):eq_primary

Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
,701	,041	,000	,621	,781

The test result variable(s): eq_primary has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

- a. Under the nonparametric assumption
- b. Null hypothesis: true area = 0.5

Coordinates of the Curve

Test Result Variable(s):eq_primary

Positive if Greater Than or Equal To ^a	Sensitivity	1 - Specificity
3,0000	1,000	1,000
4,2500	1,000	,988
4,6500	,990	,975
4,9000	,990	,963
5,0500	,980	,963
5,1500	,980	,951
5,2500	,970	,901
5,3500	,970	,889
5,4500	,960	,889
5,5500	,949	,827
5,6500	,929	,778
5,7500	,919	,728
5,8500	,919	,691
5,9500	,919	,654
6,0500	,889	,593
6,1500	,879	,568
6,2500	,848	,506
6,3500	,808	,432
6,4500	,758	,358
6,5500	,677	,333
6,6500	,606	,321
6,7500	,556	,247
6,8500	,475	,222
6,9500	,404	,185
7,0500	,263	,136
7,1500	,202	,123
7,2500	,101	,123
7,3500	,081	,111
7,4500	,061	,111
7,6500	,040	,049
7,8500	,030	,025
7,9500	,030	,000
8,1500	,010	,000
9,3000	,000	,000

The test result variable(s): eq_primary has at least one tie between the positive actual state group and the negative actual state group.

- a. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

Στους παραπάνω πίνακες πήραμε ικανοποιητικές γραφικές ενδείξεις για τη χρήση του μεγέθους του σεισμού ως μεταβλητή για την πρόβλεψη εμφάνισης τσουνάμι. Η p-value του ελέγχου για τη σημαντικότητα του εμβαδού κάτω από την καμπύλη είναι μικρότερη του 1% επομένως το μέγεθος του σεισμού δεν ταξινομεί στην τύχη τα δεδομένα. Ο δείκτης J του Youden μεγιστοποιείται για μέγεθος σεισμού 6,45 R. Για να αξιολογήσουμε την ικανότητα του συγκεκριμένου μεγέθους σεισμού να προβλέψει την εμφάνιση τσουνάμι κατασκευάσαμε μια μεταβλητή και υπολογίσαμε το ποσοστό ορθής ταξινόμησής του.

```

DATASET ACTIVATE DataSet1.
RECODE eq_primary (6.45 thru Highest=1) (ELSE=0) INTO
Youden_flag_greece.
EXECUTE.

```

flag_tsunami * Greece Crosstabulation

			Greece		Total
			0	1	
flag_tsunami	Tsunami not occurred	Count	138	29	167
		Expected Count	110,2	56,8	167,0
		% within flag_tsunami	82,6%	17,4%	100,0%
		% within Greece	68,3%	27,9%	54,6%
		% of Total	45,1%	9,5%	54,6%
	Tsunami occurred	Count	64	75	139
		Expected Count	91,8	47,2	139,0
		% within flag_tsunami	46,0%	54,0%	100,0%
		% within Greece	31,7%	72,1%	45,4%
		% of Total	20,9%	24,5%	45,4%
Total		Count	202	104	306
		Expected Count	202,0	104,0	306,0
		% within flag_tsunami	66,0%	34,0%	100,0%
		% within Greece	100,0%	100,0%	100,0%
		% of Total	66,0%	34,0%	100,0%

Το ποσοστό ορθής ταξινόμησης είναι 69,6% επομένως τουλάχιστον για την Ελλάδα και τα συγκεκριμένα δεδομένα δεν κρίνεται ικανοποιητικός δείκτης το μέγεθος του σεισμού για την πρόβλεψη εμφάνισης τσουνάμι.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική Βιβλιογραφία

1. Ηλιόπουλος Γ. (2013-2014), σημειώσεις μαθήματος “Ανάλυση διακριτών δεδομένων”, ΠΜΣ “Εφαρμοσμένης Στατιστικής”, Πανεπιστημίου Πειραιά.
2. Κούτρας Μ. (2013-2014), σημειώσεις μαθήματος “Ανάλυση Παλινδρόμησης και Διακύμανσης”, ΠΜΣ “Εφαρμοσμένης Στατιστικής”, Πανεπιστημίου Πειραιά.
3. Μπερσίμης Σ. και Σαχλάς Α. (2014), “Βιοστατιστική και στατιστικές μέθοδοι στην Επιδημιολογία για επιστήμονες υγείας”, ΠΜΣ “Εφαρμοσμένης Στατιστικής”, Πανεπιστημίου Πειραιά.
4. Μπερσίμης Σ. και Σαχλάς Α. (2014), “Στατιστική ανάλυση δεδομένων για επιστήμονες υγείας και στοιχεία μεθοδολογίας και έρευνας”, ΠΜΣ “Εφαρμοσμένης Στατιστικής”, Πανεπιστημίου Πειραιά.
5. Μελή Χ. (2012), Πτυχιακή εργασία “Χωροχρονική κατανομή της σεισμικής δραστηριότητας στον ελλαδικό χώρο με χρήση ΣΓΠ”, Χαροκόπειο Πανεπιστήμιο, Τμήμα Γεωγραφίας
6. Βασιλείου Π. (2010), Διπλωματική εργασία “Ανάπτυξη υπολογιστικών αλγορίθμων τύπου bootstrap για την επιλογή MPSS σε περιπτώσεις ανάλυσης αποτελεσματικότητας σε καθεστώς τεχνολογικής ετερογένειας”, Πανεπιστήμιο Πατρών, Τμήμα Οικονομικών Επιστημών.
7. Σώκος Θ., Σημειώσεις μαθήματος "Τεχνική Σεισμολογία", Πανεπιστήμιο Πατρών, Τμήμα Γεωλογίας.
8. Μπερσίμης Σ., Ανθοπούλου Θ. και Χαλκιάς Χ. (2007), "Ανίχνευση συστάδων έντονης επιχειρηματικότητας των γυναικών της υπαίθρου με χρήση διδιάστατων στατιστικών συναρτήσεων σάρωσης", πρακτικά 20ού Πανελληνίου Συνεδρίου Στατιστικής (ΕΣΙ).
9. Γερόδημος Φ. (2013), "Τάσεις Coulomb με δεδομένα από ιστορικούς σεισμούς", Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών.
10. Καραπατάζογλου Γ. (2013), Πτυχιακή εργασία "Η σεισμικότητα στην Ελλάδα", ΑΤΕΙ Θεσσαλονίκης, Σχολή Τεχνολογικών Εφαρμογών, Τμήμα Πολιτικών Έργων Υποδομής.

11. Χατζηδημητρίου Π. και Καρακαίσης Γ., Σημειώσεις μαθήματος "Εισαγωγή στη Σεισμολογία", ΑΠΘ, Τομέας Γεωφυσικής.
12. Σπυράκος Κ. και Τουτουδάκη Ε., "Βασικές έννοιες σεισμολογίας", Εργαστήριο Αντισεισμικής Τεχνολογίας.
13. Καλογήρου Σ. (2012), Σημειώσεις εργαστηριακού μαθήματος "Στατιστικές μέθοδοι στις Κοινωνικές Επιστήμες", Πανεπιστήμιο Πειραιώς.

Ξένη Βιβλιογραφία

1. Papadopoulos G. A. (2007), "Quantification of tsunami: The new 12-point tsunami intensity scale", International Center for Theoretical Physics.
2. Efron B. (1979), "Bootstrap methods: Another look at the Jackknife", Stanford University.
3. Yaffee R. A., Lecture notes "Robust regression modeling with Stata", New York University.
4. Cap F. (2006), "Tsunamis and Hurricanes a mathematical approach", Springer Wien New York.
5. Satake K. (2005), "Tsunamis - Case studies and recent developments", Springer.
6. Liu P. L.-F., Yeh H. & Synolakis C. (2008), "Advanced numerical models for simulating tsunami waves and runup" world scientific.
7. Soloviev S. L. and Solovieva O. N., GO C. N., Kim K. S. and Shchetnikov N. A. (2000), "Tsunamis in the Mediterranean Sea 2000 B.C.- 2000 A.D.", Springer Science & Business Media.
8. Balassanian S., Cisternas A. and Melkumyan M. (2000), "Earthquake Hazard and Seismic Risk Reduction", Springer Science & Business Media.
9. King S. A., Khalturin V. I. and Tucker B. E. (1999), "Seismic Hazard and Building Vulnerability in Post-Soviet Central Asian Republics", Springer Science & Business Media.
10. Tsuchiya Y. and Shuto N. (1995), "Tsunami Progress in Prediction, Disaster Prevention and Warning", Springer Science & Business Media.

11. Agresti A. (2007), "An introduction to categorical data analysis", John Wiley & Sons Inc.
12. Youden W.J. (1950), "Index for rating diagnostic tests", American Cancer Society.
13. Chen J. and Glaz J., "Two-dimensional discrete scan statistics", *Statistics & probability letters* 31 (1), 59-68
14. Iida K. (1956), "Earthquakes accompanied by tsunamis occurring under the sea off the islands of Japan", *J. Earth Sciences Nagoya Univ.*, 4, 1-43.
15. Iida K. (1970), "The generation of tsunamis and the focal mechanism of earthquakes. In *Tsunamis in the Pacific Ocean*", edited by W.M. Adams, East-West Center Press, Honolulu, 3 –18.
16. Iida K., Cox D.C., and Pararas-Carayannis G. (1956), "Preliminary catalog of tsunamis occurring in the Pacific Ocean", *Data Report 5, HIG-67-10*, Hawaii Inst. of Geophys. ,Univ. of Hawaii, Honolulu.
17. Imamura A. (1942), "History of Japanese tsunamis", *Kayo- No- Kagaku (Oceanography)*, 2, 74 –80 (in Japanese).
18. Imamura A. (1949), "List of tsunamis in Japan", *J. Seismol. Soc. Japan*, 2, 23-28in Japanese).
19. Iida K. (1963), "Magnitude, energy and generation mechanisms of tsunamis and a catalogue of earthquakes associated with tsunamis", *Proc. Tsunami Meeting, 10th Pacific Sci. Congress, 1961, IUGG Monograph*, 24, 7-18.
20. Soloviev S.L. and Go Ch.N. (1974), "A catalogue of tsunamis on the western shore of the Pacific Ocean", Moscow, "Nauka" Publishing House, 308p. English translation: Soloviev S.L., Go Ch.N. (1984), "A catalogue of tsunamis on the western shore of the Pacific ocean", Translation by Canada Institute for Scientific and Technical Information, National Research Council, Ottawa, Canada KIA OS2.