Ludwig-Maximilians-Universität München

Institut für Statistik

Abschlussarbeit im Rahmen des Statistischen Praktikums

Magnitudenverteilung getriggerter Erdbeben

Autorinnen: Anna Orzelek

Franziska Reichmeier

Katharina Riedlberger

Veranstaltung: Statistisches Praktikum

Abgabedatum: 02. August 2021

Projektpartner: Christian Grimm

Projektbetreuer: Dr. André Klima

**Abstract**

In der vorliegenden Arbeit werden Erdbebendaten aus Japan und Kalifornien analysiert. Genauer geht es um die Magnitudenverteilung getriggerter Erdbeben. Es wird untersucht inwiefern die Magnituden der triggernden Erdbeben einen Einfluss auf die Verteilung der von ihnen getriggerten Erdbeben haben. Dabei ist besonders von Interesse, ob starke Erdbeben auch starke Nachbeben triggern. Hierbei sollen eventuelle Unterschiede zwischen Japan und Süd-Kalifornien erfasst werden. Zudem wird die “Shortterm-Incompleteness” in der Analyse und Auswertung berücksichtigt. Diese erfolgte mit einem generalisierten additiven Modell für Lokations-, Skalen- und Formparameter (GAMLSS). ---------------------

**Inhaltsverzeichnis**

Inhaltsverzeichnis

Abbildungsverzeichnis

1. Einleitung
2. Einführung in die Thematik
3. Daten
4. Deskriptive Auswertung
5. Modelltheorie
6. Ergebnisse und Interpretation
7. Fazit

Literaturverzeichnis

Anhang

---------------------------------------------------------------------------------------

1 Einleitung

2 Einführung in die Thematik

2.1 Entstehung von Erdbeben

2.2 Begriffserläuterung bezüglich Erdbeben

3 Daten

3.1 Allgemeine Aspekte

3.2 Variablen

4 Modelltheorie

4.1 Aufbau des Modells anhand des linearen Modells

4.2 Hinzunahme einer Verteilungsannahme für die Zielvariable

4.3 Annahme eines nicht linearen Zusammenhangs

4.4 GAMLSS

5 Ergebnisse und Interpretation

5.1 Zusammenhang zwischen triggender und getriggerter Magnitude

5.1.1 Japan

5.1.2 Kalifornien

5.2 Vergleich zwischen den beiden Ländern

5.3 Umgang mit der „Shortterm-Incompleteness“

6 Fazit

**Abbildungsverzeichnis**

Abbildung1: \_\_\_\_\_\_\_\_\_\_

**1 Einleitung**

Umweltkatastrophen wie beispielsweise Erdbeben oftmals gefolgt von Tsunamis, Erdrutschen treten in einer gewissen Regelmäßigkeit auf. Es ist für Versicherungen essenziell, die eventuell entstehenden Schäden an Bausubstanzen und Menschen bereits vor Vertragsabschluss abschätzen zu können. Dafür ist es vor allem in Erdbebengebieten wichtig kalkulieren zu können, welche Auswirkungen ein Beben mit einer gewissen Stärke haben sowie den Zusammenhang zu der Stärke der Nachbeben. Durch das auslösende Erdbeben sind möglicherweise schon nicht sichtbare Schäden an Häuser und Brücken entstanden, die den Nachbeben dann nicht mehr standhalten können. Dieses Wissen ist somit unerlässlich für die Entscheidung eines Vertragsabschlusses beziehungsweise der Höhe der Versicherungsbeiträge.

Überleitung

Ziel der vorliegenden Arbeit ist es nun festzustellen, ob die Magnituden der auslösenden Erdbeben einen Einfluss auf die Verteilung der Magnituden der von ihnen getriggerten Erdbeben haben. Dabei ist von besonderem Interesse, ob starke Erdbeben auch starke Nachbeben triggern und die eventuellen Unterschiede zwischen den hier untersuchten Regionen Japan und Kalifornien zu betrachten. Insbesondere ist es von bedeutender Wichtigkeit die sogenannte „short-term incompleteness“ zu berücksichtigen.

Im Folgenden wird nun zuerst einmal die Entstehung von Erdbeben sowie wichtige Begriffe, die es in diesem Zusammenhang gibt, erläutert. Anschließend wird der bereitgestellte Datensatz vorgestellt sowie die Aufbereitung der Variablen und es folgt eine deskriptive Analyse bezüglich der Verteilung der triggernden und getriggerten Erdbeben Magnituden. Dann wird das verwendete GAMLSS Modell erklärt. Abschließend werden die Ergebnisse sowie deren Interpretation bezüglich der Fragestellung vorgestellt und in einem Fazit zusammengefasst.

1. **Einführung in die Thematik**

Im Folgenden sollen Erdbebendaten für den Projektpartner Christian Grimm ausgewertet werden. Er untersucht in seiner Doktorarbeit in Zusammenarbeit mit der Munich RE, unteranderem wie oft Erdbeben ähnlich starke Nachbeben triggern. Triggern meint in diesem Kontext, dass ein Erdbeben ein anderes auslöst, beziehungsweise aktiviert. Der Zusammenhang von sogenannten „Erdbeben-doublets“ ist interessant um Schäden, verursacht von Erdbeben, abschätzen zu können. Vor allem starke Erdbeben und ihre Triggerbeziehungen sind für die Versicherung von Interesse.

Unter Erdbeben versteht man die Erschütterung des Erdkörpers. Sie entstehen durch Plattentektoniken (oder auch an vokalischer Aktivität, Einsturz oder Absenkung unterirdischer Hohlräume, große Erdrutsche und Bergstürze sowie Sprengungen).[[1]](#footnote-1)

Ein Bild, das Karte enthält.

Automatisch generierte BeschreibungDie Theorie der endogenen Dynamik besagt, dass die äußere Erdhülle in Kontinentalplatten gegliedert ist. Diese liegen auf dem oberen Erdmantel auf und wandern auf ihm, was als Kontinentaldrift bezeichnet wird. Es gibt sieben große Lithosphärenplatten (die Nordamerikanische, Eurasische, Südamerikanische, Afrikanische, Antarktische, Australische und Pazifische Platte). Des Weiteren findet man zahlreiche kleinere Kontinentalplatten. [[2]](#footnote-2)

Abb. 1: Weltkarte mit vereinfachter Darstellung der Kontinentalplatten [[3]](#footnote-3)

Man unterscheidet in der Seismologie zwischen mehreren Arten von Plattenbewegungen: divergente, konvergente und transforme. An den Plattengrenzen findet man entweder mittelozeanische Rücken oder Tiefseerinnen. Bei der divergenten Plattenbewegung driften die benachbarten Platten auseinander. (s. Abb. 2, linkes Bild) Hier steigt oft Magma aus dem oberen Erdmantel empor und bildet eine neue Lithosphäre. Dieser Vorgang wird auch „*Seafloor Spreading*“ genannt, was meist mit intensivem unterirdischem Vulkanismus einhergeht.[[4]](#footnote-4)

Bei der konvergente Plattenbewegung driften die Platten aufeinander zu und schieben sich auf. (s. Abb. 2, mittleres Bild) Die schwerere der beiden Lithosphärenplatten taucht unter die leichtere Platte ab. Was man auch als Subduktionszone bezeichnet. In Subduktionszonen kann sich die abtauchende Platte verhaken, was folglich zu einem massiven Spannungsaufbau und letztlich zu besonders schweren Erdbeben führen kann. Diese werden auch als „Megathrust-Erdbeben“ bezeichnet. Auch hier kommt es oft zu ausgeprägtem Vulkanismus. Andernfalls kann es auch zu einer Kollision kommen, bei der eine oder beide Platten in den Randbereichen stark verformt und verdickt wird.

Die transforme Plattenbewegung meint das horizontale aneinander vorbei gleiten von Lithosphärenplatten. (s. Abb. 2, rechtes Bild) Die Bewegung verläuft dennoch nicht reibungsfrei. Die Platten verhaken sich ineinander, bis die aufgestaute Spannung zu groß wird und sich schlagartig in einem flachgründigenen Erdbeben entlädt. Auf Grund dessen sind Gebiete in der Nähe von Transformstörungen (so wie alle Regionen in der Nähe von aktiven Plattengrenzen) stark erdbebengefährdet. Ein bekanntes Beispiel für transforme Plattenbewegung ist die San-Andreas-Verwerfung in Kalifornien.[[5]](#footnote-5)



Der Datensatz für diese Projekt bezieht sich auf zwei Regionen, die an Plattengrenzen zu finden sind: Japan und Kalifornien. In der Region Japan treffen vier Platten aufeinander: die Pazifische, Eurasische, Philippinische und Nordamerikanische Platte. Hier kommt es zu einer Subduktion, bei der sich die pazifische Platte unter die eurasische Platte schiebt. In Kalifornien findet man die transforme Plattenbewegung entlang derer die Pazifische Platte an der Nordamerikanischen Platte vorbeidriftet.

Besonders an den Plattengrenzen, wenn sich die Platten in ihrer Bewegung verhaken und verkanten, bauen sich mechanische Spannungen innerhalb des Gesteins auf. Wird die Scherfestigkeit der Gesteine dann überschritten, entladen sich diese Spannungen durch ruckartige Bewegungen der Erdkruste und es kommt zum tektonischen Erdbeben.[[6]](#footnote-6)

Tiefe/Seebeben/Tsunami/vers. Erdbebentypen/ Berechnung + Unterschiede der Magnitude??

Erdbeben werden mit Hilfe eines Seismografen gemessen. Das Gerät erfasst Bodenerschütterungen von Erdbeben und anderen seismischen Wellen.[[7]](#footnote-7) Aus den gemessenen Amplituden des Seismografen bestimmt man die Magnitude. Sie bezeichnet das Maß für freigesetzte Energie, die bei einem Erdbeben entsteht. Magnitude wird auf einer Richterskala zwischen [2.0, 10.0] gemessen und gibt somit ein Maß für die Stärke eines Erdbebens.[[8]](#footnote-8) Die Skala wurde von Charles Francis Richter und Beno Gutenberg am California Institute of Technology im Jahr 1935 entwickelt. Bei einer Magnitude unter 2.0 spricht man von einem Mikrobeben, welches nicht spürbar ist. Eine Magnitude zwischen 2.0 und 5.0 wird als sehr leicht – leicht eingestuft. Diese sind teilweise spürbar, die Schäden halten sich dennoch gering. Ab einer Magnitude von 5.0 spricht man von mittelstarken Beben die (weltweit) 800-mal pro Jahr auftreten und bei stabilen Gebäuden leichte Schäden verursachen können. Bei einer Magnitude von 6.0 bis 8.0 spricht man von einem starken – sehr starken Erdbeben. Sie richten Zerstörung in Berichten bis zu einigen hunderten Kilometern an. Solche Erdbeben treten (weltweit) im Schnitt einmal im Jahr auf. Ein Erdbeben mit einer extrem großen Magnitude über 9.0 tritt durchschnittlich alle 20 Jahre auf und kann Zerstörungen in Bereichen von tausenden Kilometern verursachen. Theoretisch würde ein Erdbeben mit einer Magnitude größer gleich 10.0 eine globale Katastrophe auslösen. Dies wurde aber noch nie registriert. [[9]](#footnote-9)

* Momenten Magnitude??

Das stärkste jemals gemessene Erdbeben war das Erdbeben von Valdivia am 22. Mai 1960 in Chile mit einer Magnitude von 9.6. Es forderte 1655 Tote. Als ein besonders gravierendes wird das Tohoku-Erdbeben am 11. März 2011 in Japan mit einer Magnitude von 9.0 verzeichnet, welches auch in dem Datensatz enthalten ist. Insgesamt 18.500 Menschen wurden Opfer des schweren Erdbebens und nachfolgenden Tsunamis. Es entstand ein Schaden von etwa 296 Milliarden Euro.[[10]](#footnote-10)

In diesem Zusammenhang ist auch das Gutenberg-Richter-Gesetz zu erwähnen. Es beschreibt die Beziehung zwischen der Stärke und der Gesamtzahl von Erdbeben in einer bestimmten Region und einem Zeitraum von mindestens dieser Stärke. N gibt die Anzahl der Erdbebenereignisse mit N M an.

Dies bedeutet, dass es bei einem Erdbeben mit Magnitude 5,0 zehnmal so viele Beben der Stärke 4,0 und 100-mal so viele Beben der Stärke 3,0 auftreten. Diese Beziehung zwischen Magnitude und Häufigkeit des Auftretens ist bemerkenswert verbreitet, obwohl die Werte von a und b erheblich von Region zu Region variieren können sowie im Laufe der Zeit.[[11]](#footnote-11)

Da Erdbeben meistens nicht aus einer einzelnen Erschütterung bestehen, sondern weitere nach sich ziehen, spricht man in diesem Zusammenhang von Erdbebencluster. Hierbei stehen mehrere Erdbeben in einem regionalen und zeitlichen Zusammenhang. In Abbildung 3 stellen die einzelnen Boxen schematisch Erdbeben mit der jeweiligen Magnitude dar. Die Pfeile geben die Triggerbeziehungen an. Das heißt, wenn ein Erdbeben getriggert wurde, zeigt ein Pfeil, ausgehend vom triggernden Beben, auf dessen Box. Das Beben mit Magnitude 8.0 beispielweise wurde von dem Beben der Magnitude 4.0 getriggert und triggert selbst ein weiteres Beben mit Magnitude 3.0. Dabei kann es auch vorkommen, dass ein Beben mehrere Nachbeben triggert.

Auf der linken Seite des Baumes steht eine Ellipse stellvertretend für die Shortterm-Incompleteness Phase. Sie bezeichnet die Phase in dem Erdbeben, die durch ein zeitlich vorangegangenes starkes Erdbeben nicht vollständig erfasst werden konnten. Im Falle der Abbildung 3. war das triggernde Beben mit der Magnitude so stark, dass seine Folgebeben nicht vollständig erfasst werden konnte. Das starke Beben übertönt die Messungen der Seismografen. Insbesondere Erdbeben mit geringer Magnitude können in diesem Zeitraum nicht erfasst werden. Abhängig von der Stärke des Bebens, welches die Shortterm-Incompleteness Phase auslöst, ist diese Phase unterschiedlich lang. Je näher man sich zeitlich an dem auslösenden Beben befindet, umso schwerer können niedrige Beben erfasst werden und umso stärker ist somit die Verzerrung der Messungen.[[12]](#footnote-12)

Außerdem lassen sich auch so genannte Single Events verzeichnen, welche Erdbeben meinen, die nicht getriggert wurden und keine weiteren Erdbeben triggern.

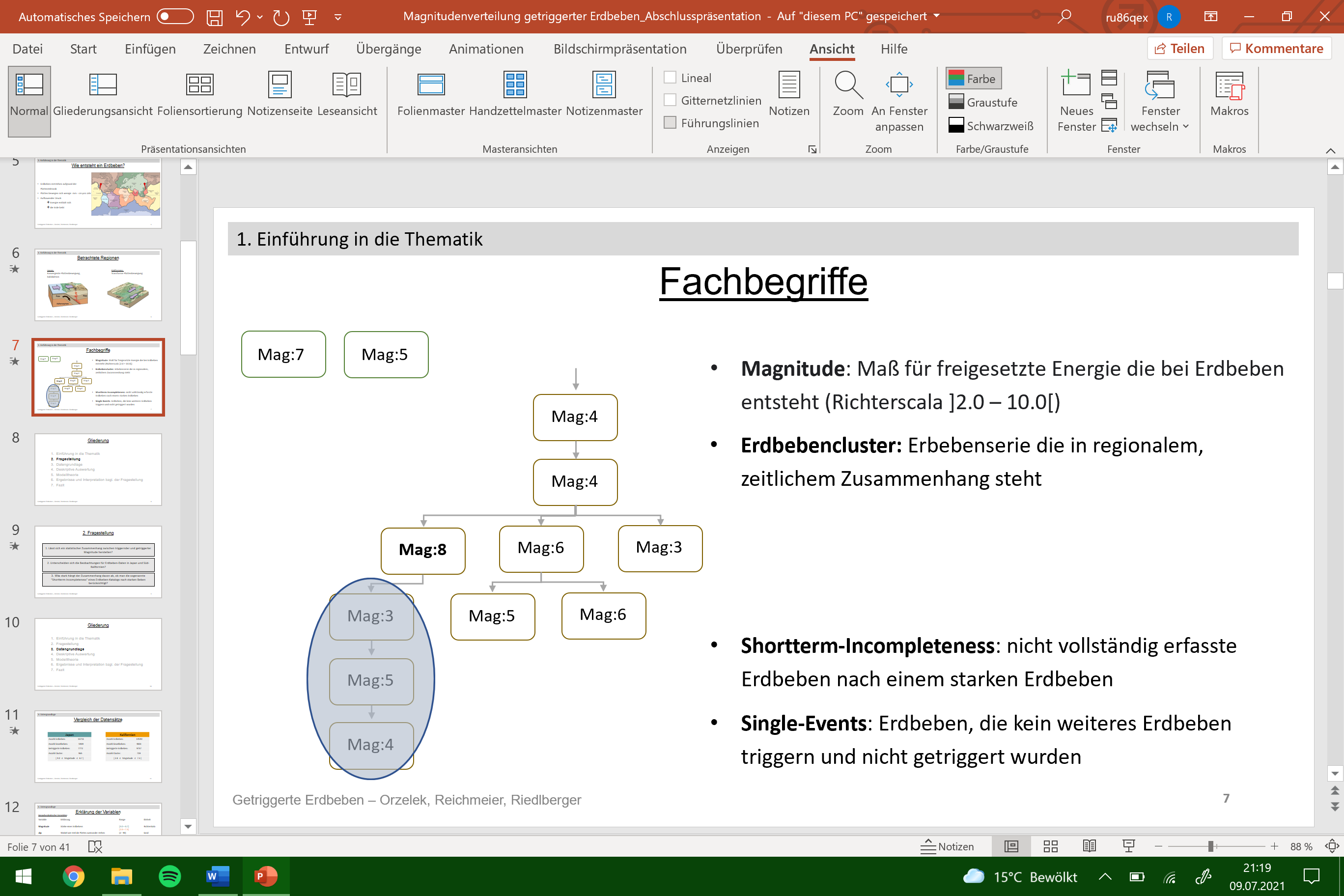


Abb. 3: Schematische Skizze eines Erdbebenclusters. Die Einzelnen Boxen stellen Erdbeben mit jeweiliger Magnitude dar. Pfeile zwischen den Boxen geben Triggerbeziehungen an. Ellipse über dem linken Ast soll stellvertretend die Shortterm Incompleteness Phase darstellen. Zudem finden sich oben links zwei Einzelbeben.

Der Großteil der aufgezeichneten Erdbeben ist zu schwach, um vom Menschen wahrgenommen zu werden. Dennoch richten vor allem starke Erdbeben erhebliche Schäden an. Beispielsweise werden oft Bauten vernichtet zudem werden Tsunamis, Lawinen, Steinschläge, Bergstürze und Erdrutsche ausgelöst, wobei Menschen verletzt oder getötet werden können. Die einhergehenden Zerstörungen der Erdoberfläche sind auf die „Oberflächenwellen“ zurückzuführen, die sich an der Erdoberfläche ausbreiten und eine elliptische Bodenbewegung auslösen. Das Ausmaß der Schäden hängt aber von der Stärke und Dauer des Bebens ab, sowie der Bevölkerungsdichte, und Anzahl und Größe der Bauwerke in der betroffenen Region.[[13]](#footnote-13)

**2 Datengrundlagen**

**2.1 Allgemeine Aspekte**

Die Daten wurden in Japan von Januar 1997 bis Oktober 2020 erfasst und in Kalifornien im Zeitraum Januar 1981 bis Dezember 2019. Es handelt sich bei den betrachteten Daten um Vollerhebungen.

Während dieser ca. 24 Jahren wurden in Japan 13711 Erdbeben gemessen, davon waren 5939 Einzelbeben, die in der weiteren Analyse entnommen wurden, da sie für die Analyse nicht relevant sind. Folglich wurden in der Auswertung 7772 Erdbeben betrachtet. Zudem befanden sich 965 Cluster in diesem Teildatensatz.

In Kalifornien haben während der 38 Jahre nachweislich 14540 Erdbeben stattgefunden. Hier waren für die folgende Analyse 4803 Erdbeben irrelevant, da sie Einzelbeben waren. Folglich wurden 9737 Erdbeben weiterhin betrachtet, die in 719 Clustern auftraten.

**2.2 Variablen**

Wie man der Fragestellung entnehmen kann, soll primär der Fokus auf den Einfluss der Magnitude des triggernden Bebens auf die Magnitude des getriggerten Beben gelegt werden.

Mithilfe eines Algorithmus, welcher mit der räumlichen und zeitlichen Distanz der Erdbeben gearbeitet hat, wurde bereits vorab vom Projektpartner bestimmt, ob das Erdbeben getriggert wurde und wenn ja von welchem.

Wie zuvor bereits erwähnt wurde, wird die Magnitude in einem Intervall von 2.0 bis 10.0 erfasst. Bei den Daten aus Japan wurden die Erdbeben ab einer Magnitude von vier erfasst. Erdbeben mit einer niedrigeren Magnitude wurden dem Datensatz bereits vorab entnommen, da man aufgrund des bereits erwähnten Gutenberg-Richter-Gesetzes feststellen konnte, das der Erdbebenkatalog nicht vollständig vorlag. Das stärkste Beben was in dieser Region gemessen werden konnte, war das Tohoku-Erdbeben am 11.03.2011 mit einer Magnitude von 8.7.

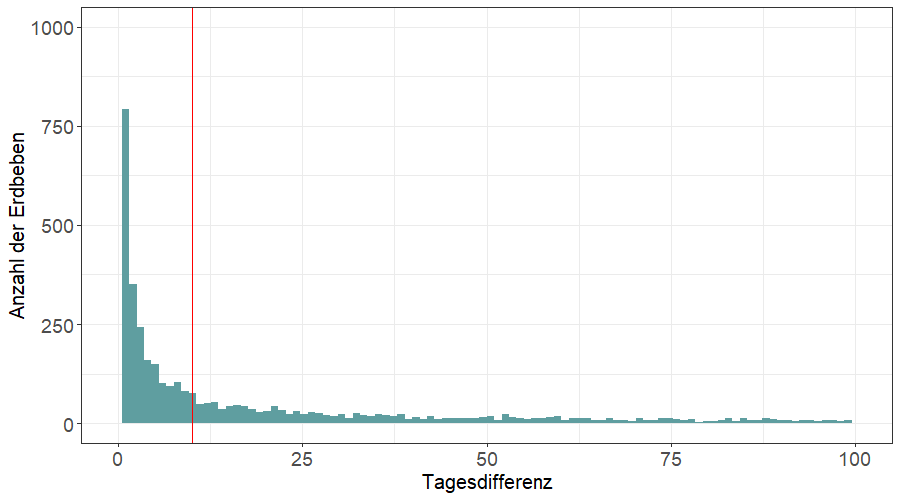
Dies gilt analog für Kalifornien. Aufgrund besserer Messtechnik ist der Erdbebenkatalog erst ab einer Magnitude kleiner als 2.8 unvollständig. Allerdings hatten die stärksten Beben eine niedrigere Magnitude (7.3) verglichen zu Japan. Dies kann auf die bereits erwähnten verschiedenen Plattenbewegungen in Japan und Kalifornien zurückgeführt werden.

In die Berechnung des Modells flossen noch weitere physikalischen Variablen mit ein, die nun kurz erläutert werden.

Der „dip“ gibt an in welchem Steilwinkel die Platten zueinanderstehen. Im Gegensatz dazu sagt der „rake“ aus, in welchem Winkel die Platten sich relativ auf- und abschieben. Der Wärmefluss aus dem Erdinneren an die Oberfläche während eines Erdbebens wird „heat flow“ genannt. „depth“ gibt die Tiefe unter der Oberfläche an, an der das Erdbeben stattgefunden hat. „mantle thickness“ ist die Erdmanteldicke und analog dazu floss noch die „crustal thickness“, also die Erdkrustendicke, mit ein. „strain rate“ gibt an, was die Deformationsrate ist, die durch die resultierenden Plattenbewegungen entstanden ist.

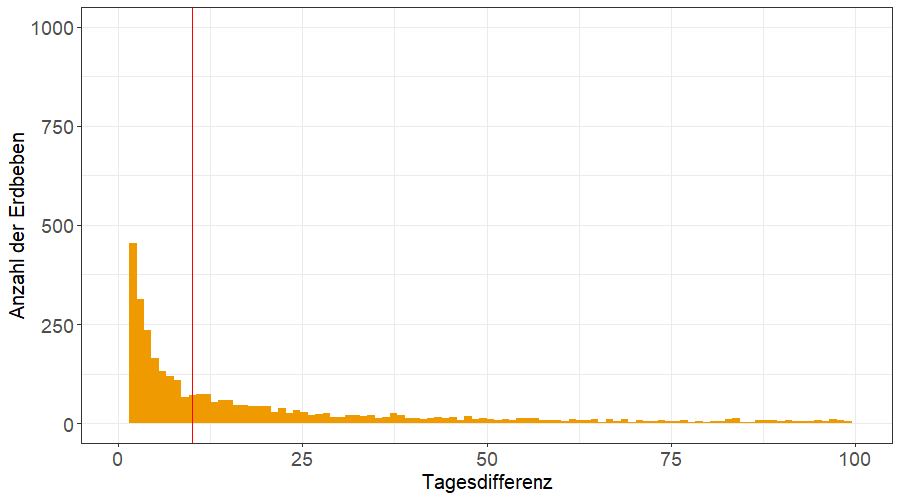
Eine weitere Variable ist die sogenannte „Completeness Magnitude“, die keinen direkten physikalischen Bezug hat, sondern von unserem Projektpartner konstruiert wurde. Sie sagt aus, welche die aktuell niedrigste messbare Magnitude war, als das Erdbeben stattgefunden hat. Wenn es möglich war alle Beben vollständig zu erfassen, so besitzt das Beben in Japan eine „Completeness Magnitude“ von 4.0 bzw. in Kalifornien von 2.8. Daraus ergibt sich somit die Blindheitsphase, auch Shortterm-Incompleteness genannt. Wenn alle Beben erfassbar sind, hat das gemessene Erdbeben nicht während der Blindheitsphase stattgefunden. Ist die „Completeness Magnitude“ größer als vier in Japan bzw. größer als 2.8 in Kalifornien, ist davon auszugehen, dass das Erdbeben, während einer Shortterm-Incompleteness Phase stattgefunden hat. Man kann davon ausgehen, dass umso höher die „Completeness Magnitude“ ist, desto größer ist die Verzerrung der unvollständig erfassten Erdbeben, die es zu dem Zeitpunkt gab. Somit ergab sich eine Variable, die angibt, ob ein Beben während einer Blindheitsphase stattgefunden hat oder nicht.

Bei der Zeitdifferenz zwischen triggernden und getriggertem Beben gab es zum Teil große Differenzen von ungefähr 7900 Tagen in Japan. Bezüglich des zeitlichen Abstands einer Trigger-Relation gibt es nun zwei Positionen in der Seismologie. Die eine Ansicht besagt, dass die Relationen einen großen zeitlichen Abstand haben können, von 20 Jahren und mehr. Das birgt die Gefahr, dass Erdbeben als getriggert gelten, die es aber nicht sind. Anderseits gibt es die Theorie, die besagt, dass die Zeitdifferenz nicht größer als 24 Stunden sein kann, folglich kann es vorkommen, dass der Erdbebenkatalog dann unvollständig ist. Erdbeben mit einer großen Zeitdifferenz haben die Effekte verzerrt und damit diese die Effekte nicht dominieren, wurde die Entscheidung getroffen einen Schwellenwert festzulegen. Betrachtet man nun die Abbildung X, so sieht man die Zeitdifferenz bis zu 100 Tagen in Japan (das Maximum liegt viel höher, aber die Daten dünnen sich nach 100 Tagen so sehr aus, dass man nichts mehr sieht). Man erkennt nun, dass ein Großteil der Beobachtungen eine Zeitdifferenz von kleiner gleich 10 Tagen (analog auch für Kalifornien) haben und somit fiel die Entscheidung bezüglich des Schwellenwertes in Absprache mit dem Projektpartner. Das heißt, dass alle Zeitdifferenz größer als 10 Tage auf diesen Wert gesetzt wurden.[[14]](#footnote-14) So veränderte man in Japan rund 40 % der beobachteten Zeitdifferenzen und in Kalifornien waren es gut 30 %.



Die Histogramme für Zeitdifferenz in Kalifornien befinden sich im Anhang unter \_\_\_\_, sowie genauere Informationen bezüglich der Variablen im Anhang unter \_\_\_\_.

Anhang



**3 Deskriptive Auswertung**

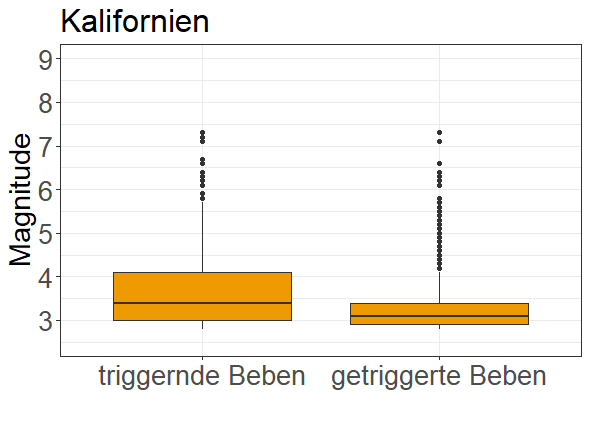
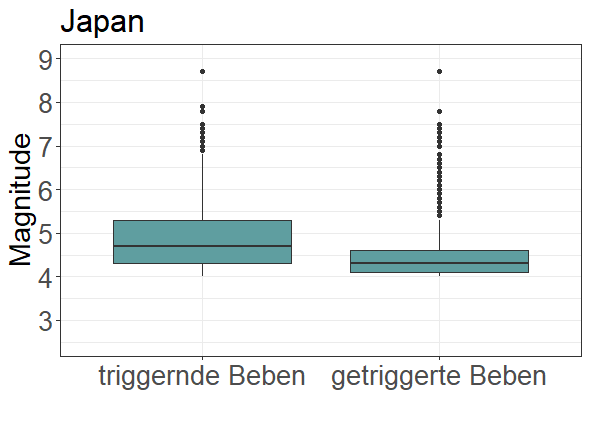


Abbildung 1:

Im Folgenden werden alle Graphiken bezüglich Japan blau beziehungsweise Kalifornien orange gehalten (s. Abb \_\_). Diese Abbildung dient dazu, einen Überblick über die Magnitudenverteilung der triggernden beziehungsweise getriggerten Beben zu erhalten. Die Range der Magnituden Verteilungen der triggerenden und getriggerten Erdbeben in Japan geht von 4.0 bis 8.7. Betrachten wir nun die Magnituden Verteilung der triggernden Erdbeben, also der Einflussgröße in Japan. Das 0.25-Quantil liegt bei 4.3 und der Median bei 5.3. Es ergibt sich ein Interquartilsabstand von einer Magnitudenstufe. Vergleicht man dies nun mit den getriggerten Erdbeben, fällt auf, dass das 0.25-Quantil bei 4.1 liegt. Auch der Median ist niedriger. Folglich ergibt sich ein Interquartilsabstand von einer halben Magnitudenstufe. Bei Kalifornien ergibt sich ein ähnliches Bild. Das 0.25-Quantil sowie der Median der triggernden Beben ist etwas höher als bei den getriggerten Erdbeben. Auch der Interquartilsabstand der triggernden Beben ist gut doppelt so groß wie bei den getriggerten Beben. Allerdings geht die Range der Magnituden der triggerenden sowie getriggerten Beben von 2.8 bis 7.3.

Abschließend ist zu sagen, dass alle vier Verteilung sehr linkssteil beziehungsweise rechtsschief sind.

Modelltheorie

LM

Das allgemeine Ziel der Regression besteht darin, eine Beziehung zwischen der interessierenden bzw. der Zielvariable Y und einer oder mehreren Einflussvariablen x1, …, xp herzustellen, wobei bei der linearen Regression von einem linearen Zusammenhang ausgegangen wird. Die zielvariable sowie die abhängigen Variablen sind metrische Größen. Daraus ergibt sich für eine Stichprobe von Umfang i = 1, 2, …, n das lineare Modell als:

Yi = β0 + β1 \* xi1 + β2 \* xi2 + … + βp \* xip + εi = **xtiβ** + εi

Mit xi = (1, xi1, xi2, ... , xip)talsPrädiktorvektor und β als Parametervektor.[[15]](#footnote-15) Dabei gelten folgende Annahmen. Die Beobachtungen der Zielvariable Yi sind unabhängig. Außerdem haben die Resiuduen εi E(εi)=0 und Var(εi)=σ2. Sind die Residuen zusätzlich normalverteilt, spricht man von klassischer Normalregression.[[16]](#footnote-16) Man schätzt somit den bedingten Erwartungswert von Y bei gegebenen **xti** und das Modell liefert eine Prognose der Zielvariablen.

Im hiesigen Fall würde die Modellgleichung folgendermaßen aussehen:

Getriggerte Magnitudei = β0 + β1 \* xi triggernde Magnitude + β2 \* xi heat flow + β3 \* xi strain rate +

β4 \* xi dip + β5 \* xi depth + β6 \* xi rake + β7 \* xi crustal thickness + β8 \* xi zeitdifferenz +

β9 \* xi completeness Magnitude + β10 \* xi mantle thickness

Da primär der Zusammenhang zwischen triggernder und getriggerter Magnitude betrachtet werden soll, ist die getriggerte Magnitude die Zielvariable und die triggernde Magnitude ist als Einflussvariable im Modell aufgenommen. Indem die anderen Variablen auch im Modell enthalten sind, wird deren Effekt auf die getriggerte Magnitude kontrolliert. Die Schätzung der unbekannten Koeffizienten β erfolgt anhand der Kleinste-Quadrate-Methode (KQ-Methode). Diese beruht auf der Minimierung der Summe der quadrierten Abweichungen:[[17]](#footnote-17)

Man würde folglich ein Modell erhalten, welches eine erste Prognose für den Erwartungswert der getriggerten Magnitude liefert. Bei Betrachtung der Verteilung der Zielvariable fällt auf, dass diese Exponentialverteilt ist (vgl Abb. 1). Gezeigt ist hier ein Histogramm, welches die Häufigkeitsverteilung der getriggerten Magnituden für Japan widerspiegelt. Zu beachten ist, dass die Magnituden so transformiert wurden, dass sie bei 0 beginnen, indem alle Werte mit 4 subtrahiert wurden. Dadurch liegt das Minimum der Magnituden bei 4, wodurch wird der Vergleich mit der Exponentialverteilung vereinfacht.

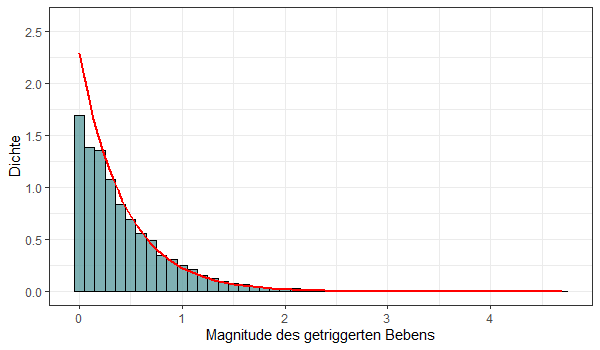


Abb. 1: Vergleich der Verteilung der getriggerten Magnituden für Japan mit der Dichte der Exponentialverteilung mit λ = 2.28 (rote Linie)

Diese Verteilung folgt dem Gutenberg-Richter-Gesetz: beobachtet man beispielsweise ein Beben der Magnitude6, treten 10-mal so viele Beben der Magnitude 5.0 auf und 100-Mal so viele der Magnitude 4.0.[[18]](#footnote-18) Man hat somit einen exponentiellen Zusammenhang. Diese Annahme wird durch die Dichte der Exponentialverteilung mit Parameter λ = 2.28 bestätigt, die hier durch eine rote Linie dargestellt wird. Der Parameter λ lässt sich folgendermaßen schätzen:

Erstellt man eine identische Graphik für Kalifornien, zeichnet sich ein vergleichbares Bild ab. Aus verschiedenen Gründen, die später genauer erläutert werden, wird im Weiteren eine Gammaverteilung angenommen. Dies ist möglich, da die Exponentialverteilung ein Spezialfall der Gammaverteilung ist, bei dem der erste Parameter auf eins gesetzt wird. Die Gammaverteilung Ga(a, b) ist definiert durch die Dichte[[19]](#footnote-19):

für alle a, b > 0 (1)

Folglich ergeben sich Erwartungswert und Varianz zu:

und (2)

Die Daten erfüllen durch ihren Definitionsbereich auf den positiven Reellen zahlen und ihre Rechtsschiefe außerdem die Eigenschaften der Gammaverteilung. Möchte man diese Information in das Modell miteinfließen lassen, betrachtet man kann einfaches lineares Modell mehr, sondern ein generalisiertes lineares Modell (GLM). Dieses erlaubt es, eine Verteilungsannahme für die Zielvariable zu treffen. Die angenommene Verteilung unterliegt hierbei einer einparametrigen linearen Exponentialfamilie. Die Überprüfung dieser Voraussetzung ist hier jedoch nicht notwendig, da das finale Modell diese Anforderung nicht hat. Ein GLM zeichnet sich durch den linearen Prädiktor ηi = xtiβ und eine stetige, zwei Mal differenzierbare Linkfunktion aus. Diese verbindet den linearen Prädiktor mit dem Erwartungswert und hängt von der getroffenen Verteilungsannahme ab.[[20]](#footnote-20) In dem Fall der Gammaverteilung ergibt sich ein log-Link. Der Zusammenhang sieht somit wie folgt aus:

g(E(Y)) = xtiβ ⬄ E(Y) = g-1(xtiβ)

wobei g der Logarithmus und g-1 somit die Exponentialfunktion ist. Die Modellgleichung bleibt dadurch unverändert, lediglich die dahintersteckenden Annahmen passen sich an: Für die Zielvariable Y wird eine Gammaverteilung angenommen, wodurch der log-Link bei der Auswertung der Effekte berücksichtigt werden muss.

Ziel ist es nun, die Form des Einflusses der Kovariablen besser abzubilden, da eine lineare Modellierung eventuell nicht ausreichend ist. Die Idee hierbei ist es, die Effekte nicht mehr rein parametrisch zu schätzen, sondern über eine Summe glatter Polynomialfunktionen. Dafür wird der Wertebereich der Einflussvariable unterteilt und für jeden Abschnitt ein Polynom gerechnet. Die Grenzen der Bereiche werden als Knoten bezeichnet. Man spricht dann von einem generalisierten additiven Modell (GAM). Man erhält dann folgende Modellgleichung:

g(µi) = f1(zi1) + … + fq(ziq) + xtiβ

Hierbei sind z1, …, zq stetige Variablen, denn nur diese können glatt geschätzt werden. Alle vorigen Modellannahmen des GLM gelten weiterhin.

Es gibt verschiedene Ansätze für die Anwendung solcher Polynomialfunktionen. Einer davon sind B-Splines. Diese stellen den Anspruch, dass die Enden der Polynome kontinuierlich verlaufen und die Funktion somit insgesamt stetig ist. [[21]](#footnote-21)Hier werden die einzelnen Polynome als Basisfunktionen bezeichnet. Deren Summe ergibt sich nun folgendermaßen:

d bezeichnet die Anzahl der Basisfunktionen, welche Anhand der Zahl der inneren Knoten festgelegt wird. l gibt den Polynomgrad der Basisfunktion an. In der Praxis werden größtenteils kubische Polynome, also Polynome vom Grad 3 verwendet, da sie eine ausreichende Flexibilität bieten ohne Überanpassung an die Daten (genannt Overfitting). γ stellt den Vektor der Gewichte der einzelnen Basisfunktionen dar. Diese Summe kann somit in Form eines linearen Modells dargestellt werden und eine Schätzung des Gewichtsvektors ist über die Kleinste-Quadrate-Methode möglich. Das Problem hierbei ist, dass die Flexibilität des Glätters stark von der Anzahl innerer Knoten abhängt. Wählt man jedoch eine zu hohe Anzahl an Knoten, droht Overfitting. Um dem entgegenzuwirken, bieten die sogenannten penalisierten Splines (P-Splines) eine Möglichkeit zur automatisierten Wahl der Komplexität des Glätters. Dafür wird im ersten Schritt die zu schätzende Funktion f(x) durch Polynom-Splines mit einer sehr hohen Knotenanzahl approximiert, wodurch man eine große Flexibilität erreicht. Anschließend wird eine zu raue Schätzung durch einen Strafterm penalisiert. Man versucht dadurch eine zu starke Anpassung an die Daten zu vermeiden und gleichzeitig die Glättung des Effekts beizubehalten. Dafür wird der Gewichtsvektor γ hier über die Minimierung eines penalisierten KQ- Kriterium geschätzt:

P(γ) ist der Strafterm und der Glättungsparameter λ steuert hier den oben genannten Trade-Off: lässt man λ gegen unendlich gehen, so wird die Schätzung des Gewichtsvektors von dem Strafterm bestimmt und die Funktion f(x) passt sich zu stark an die Daten an. Hält man λ hingegen klein, so geht der Strafterm nur mit sehr geringem Gewicht in die Schätzung ein und es ergibt sich eine nahe am KQ-Schätzer gelegene Schätzung von γ. Der Vorteil der penalisierten Splines liegt somit darin, dass die Glättung anhand des Parameters λ gesteuert werden kann und unabhängig von Anzahl und Position der Knoten erfolgt.[[22]](#footnote-22) Es existieren zudem noch weitere Spline-Arten, insbesondere zyklische Splines, welche bei dem finalen Modell Anwendung finden. Sie beruhen auf der Annahme, dass Minimum und Maximum der Variable dieselbe inhaltliche Bedeutung haben und deren Effekte somit identische geschätzt werden.[[23]](#footnote-23)

Ziel der Auswertung ist es nun, die Verteilung der getriggerten Magnituden zu betrachten. Deshalb ist es notwendig nicht ausschließlich den Mittelwert zu betrachten, sondern auch weitere Verteilungsparameter in die Analyse miteinzubeziehen. Dementsprechend fiel die Entscheidung auf ein GAMLSS, ein generalisiertes additives Modell für Lage, Skalen- und Formparameter, welches nicht nur Mittelwertsregression, sondern eine flexible Modellierung der Zielvariable mithilfe weiterer Verteilungsparameter erlaubt. Es handelt sich hierbei um eine semiparametrische Regression: man hat die Möglichkeit die Effekte mithilfe von nicht-parametrischen Splinefunktionen glatt zu schätzen, jedoch wird, wie beim GLM eine parametrische Verteilungsannahme benötigt. Wie zuvor beschrieben, wird im Folgenden eine Gammaverteilung für unsere Zielvariable, die getriggerte Magnitude angenommen. Zusätzlich zu den bereits erwähnten Gründen, bietet diese Verteilung den Vorteil zwei Verteilungsparameter schätzen zu können. Es ergibt sich dadurch folgende Parametrisierung der Dichte Ga(σ - 2, (σ2µ)-1) aus (1):

wobei µ und σ die Bezeichnungen für die zu schätzenden Parameter des GAMLSS sind. Es gilt außerdem µ, σ > 0. Dadurch erhält man folgenden Erwartungswert und Varianz (siehe (2)):

und = σ2µ2

Letztendlich ergibt sich das finale Modell, jeweils für die Parameter µ und σ folgendermaßen:

Log(µ) = β0 + f1(triggernde Magnitude) + f2(heat Flow) + f3(strain rate) + f4(dip) + f5(depth) +

f6(rake) + f7(crustal thickness) + f8(Zeitdifferenz) + f9(completeness Magnitude) +

f10(mantle thickness)

Log(σ) = β0 + f1(triggernde Magnitude) + f2(heat Flow) + f3(strain rate) + f4(dip) + f5(depth) +

f6(rake) + f7(crustal thickness) + f8(Zeitdifferenz) + f9(completeness Magnitude) +

f10(mantle thickness)

wobei f1, …, f10 wie zuvor beschrieben für nicht-parametrische Splinefunktionen stehen. Hier wurden penalisierte Splines verwendet, bis auf f6 welche für zyklische Splines steht. Um die daraus geschätzten Werte sinnvoll als Effekte auf den Erwartungswert beziehungsweise die Varianz zu interpretieren, sind folgende Transformationen notwendig:

Effekt auf Erwartungswert = exp(µ) Effekt auf Varianz = exp(σ)2 \* exp(µ)2

Zur Erinnerung, das Anwenden der Exponentialfunktion ist aufgrund des log-Links erforderlich. Dadurch lassen sich die erhaltenen Effekte als multiplikative Effekte auf den Erwartungswert, beziehungsweise die Varianz der triggernden Magnitude interpretieren.

IV. Ergebnisse

1. Zusammenhang zwischen triggernder und getriggerter Magnitude

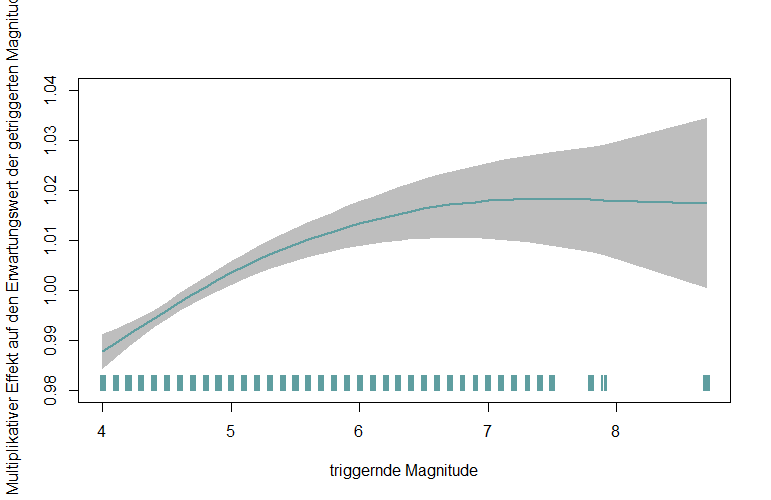
Für die erste Fragestellung, in der der statistische Zusammenhang zwischen triggernder und getriggerter Magnitude untersucht werden soll, betrachten wir im Folgenden ein Modell, das nach zuvor aufgestellter Theorie entstanden ist. Um also eine Aussage über die Verteilung der Magnituden der getriggerten Beben treffen zu können, wird jeweils ein Modell für Japan und für Kalifornien betrachtet. Diese werden für die Verteilungsparameter µ und σ gerechnet, woraus sich, wie im vorigen Teil beschrieben, der Effekt auf den Erwartungswert und die Varianz herleiten lässt. Die Modelle haben als Zielvariable die getriggerte Magnitude der jeweiligen Region und als Einflussgrößen die triggernde Magnitude des jeweiligen Bebens und die dazugehörigen Kovariablen.

1a. Japan

Das folgende Modell wird für die Auswertung der Japan-Daten verwendet:

Getriggerte Magnitude ~ pb(triggernde Magnitude) + pb(heat Flow) + pb(strain Rate) + pb(dip) + pb(depth) + cy(rake) + pb(mantle Thickness) + pb(crustalThickness) + pb(completeness Magnitude) + pb(Zeitdifferenz)

wobei “pb” für penalisierte und “cy” für zyklische Splines steht. Die Gleichung ist für die beiden betrachteten Parameter identisch.



Triggernde Magnitude

Multiplikativer Effekt auf den Erwartungswert

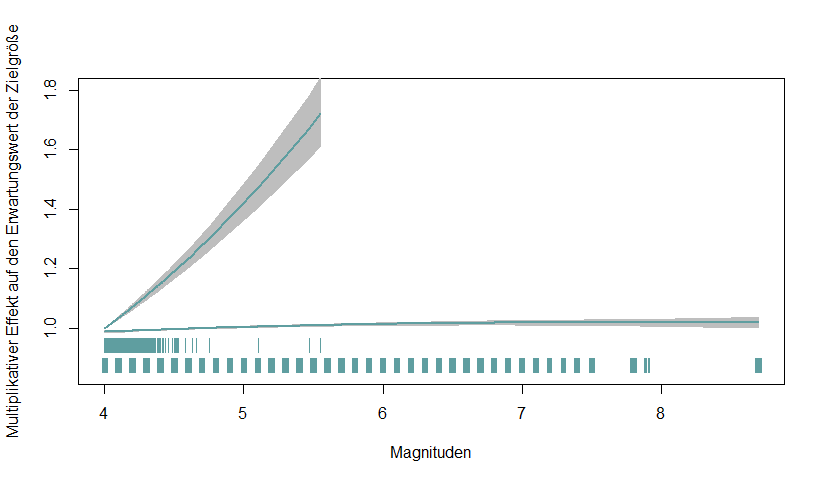
der getriggerten Magnitude

Abb A: Multiplikativer Effekt der triggernden Magnitude auf den Erwartungswert der getriggerten Magnitude für Japan

Als erstes betrachten wir in Abbildung A den multiplikativen Effekt der triggernden Magnitude auf den Erwartungswert der getriggerten Magnitude in Japan, da die Fragestellung primär auf diesen Zusammenhang abzielt. Zu erkennen ist ein konstant ansteigender Effekt, der ab einer triggernden Magnitude von 6.5 abflacht. Zu beachten ist jedoch, dass dieser Effekt sich in einem sehr kleinen Bereich bewegt, der Effekt steigt über den kompletten Magnitudenbereich lediglich um den Faktor 0.03 an. Es fällt außerdem auf, dass sich die Daten ab einer Magnitude von 7.5 ausdünnen und das 95%-Konfidenzintervall des geschätzten Effekts zunimmt. Das Ausdünnen der Daten lässt sich an dem Balken im unteren Bereich der Graphik ablesen. Dieser stellt die Verteilung der Datenpunkte für die jeweilige Variable dar. Bei der Prüfung der anderen im Modell aufgenommen Effekte fällt auf, dass sie sich in einem ähnlich kleinen, teilweise noch kleineren Bereich bewegen. Einzig die completeness Magnitude (Abb. B) weist einen größeren Effekt auf. Dies ist auch in den weiteren Modellen der Fall, weshalb der Effekt der triggernden Magnitude im weiteren Verlauf stets mit dem Effekt der completeness Magnitude verglichen wird.

Magnitude

Multiplikativen Effekt auf den Erwartungswert der getriggerten Magnitude



Completeness Magnitude

Triggernde Magnitude

ABB B: Multiplikative Effekte der triggernden sowie der completeness Magnitude auf den Erwartungswert der getriggerten Magnitude für Japan

Hier ist zu erkennen, dass die completeness Magnitude einen deutlich stärkeren Effekt auf den Erwartungswert der getriggerten Magnitude aufweist als die triggernde Magnitude. Man beobachtet annähernd eine Verdoppelung des Erwartungswerts über den Gesamtbereich der completeness Magnitude. Das bedeutet, dass die, durch die Unvollständigkeit des Erbebenkatalogs während einer Blindheitsphase verursachte Verzerrung der Daten, den größten Effekt auf den Erwartungswert der getriggerten Magnitude hat. Im Vergleich dazu, ist der Effekt der triggernden Magnitude äußerst gering. Beobachtet man beispielsweise eine completeness Magnitude von 5.0, so ist die getriggerte Magnitude im Mittel um den Faktor 1.4 größer als bei einer completeness Magnitude von 4.0.

Kommen wir nun zu dem Effekt auf die Varianz. Hierzu beginnen wir erneut mit dem Effekt der triggernden Magnitude, dargestellt in Abb. C.

Ein Bild, das Text, Messstab enthält.

Automatisch generierte Beschreibung

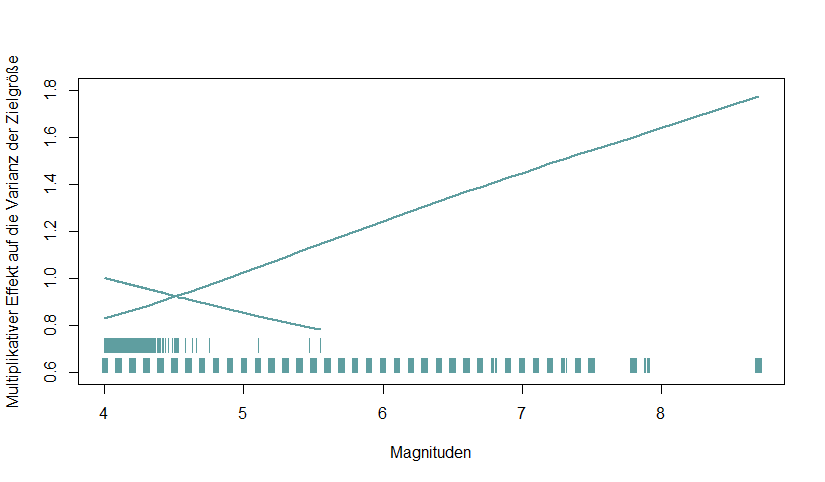
Multiplikativer Effekt auf die

Varianz der getriggerten Magnitude

Triggernde Magnitude

Abb. C: Multiplikativer Effekt der triggernden Magnitude auf die Varianz der getriggerten Magnitude für Japan

Zu beobachten ist ein linear ansteigender Effekt. Dieser spiegelt eine Verdopplung der Varianz über den gesamten Bereich wider. Hier ist die Trunkierung der Daten nach unten auf eine Magnitude von 4 zu beachten, woraus sich mit steigender triggernder Magnitude eine erhöhte Wahrscheinlichkeit für starke Nachbeben ableiten lässt. Da die Daten nach unten begrenzt sind, können die Magnituden bei einer Erhöhung der Varianz lediglich nach oben streuen. Zudemfällt auf, dass bei den Effekten auf die Varianz keine 95%- Konfidenzintervalle für die Effekte abgebildet sind. Dies liegt daran, dass nicht mit Sicherheit von einer unabhängigen Berechnung von µ und σ ausgegangen werden kann und eine Darstellung der Konfidenzintervalle somit womöglich nicht korrekt wäre. Als Alternative zu Abschätzung der Unsicherheit, bietet es sich an die Verteilung der Datenpunkte zu betrachten.



Multiplikativer Effekt auf die

Varianz der getriggerten Magnitude

Magnitude

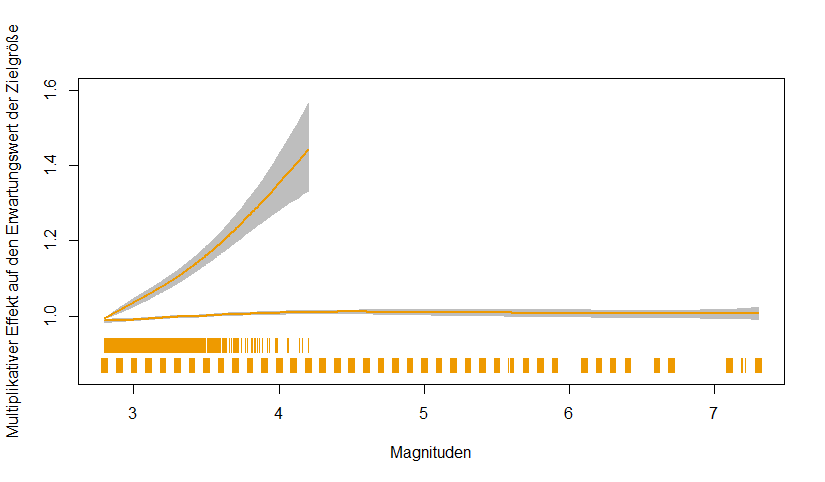
Abb. D: Multiplikative Effekte der triggernden sowie der completeness Magnitude auf die Varianz der getriggerten Magnitude für Japan

Vergleichen wir diesen Effekt nun erneut mit dem der completeness Magnitude (Abb. D), stellt man fest, dass die triggernde Magnitude einen weitaus größeren Effekt auf die Magnitude der getriggerten Beben hat als die completeness Magnitude. Deren Effekt ist linear absteigend von 1.0 bis 0.8. Der negative Effekt lässt sich dadurch erklären, dass während einer Shortterm-Incompleteness Phase, die hier durch eine completeness Magnitude ungleich 4 repräsentiert wird, Beben aus einem eher niedrigen Magnitudenbereich nicht erfasst werden. Man beobachtet somit nur noch einen höheren, eingeschränkten Magnitudenbereich und die Varianz der Daten nimmt mit ansteigender completeness Magnitude ab. Letztendlich lässt sich ein statistischer Zusammenhang zwischen triggernder und getriggerter Magnitude in Japan feststellen, da die Magnitude der triggernden Beben einen starken Einfluss auf die Varianz der Magnitude der getriggerten beben aufweist.

1b. Kalifornien

Im folgenden Abschnitt werden wir uns mit derselben Fragestellung wie zuvor beschäftigen, dieses Mal jedoch für die Erdbebendaten in Kalifornien. Da die betrachtete Problematik identisch bleibt, wird auch das Gleiche Modell benutzt, mit der Ausnahme, dass mantle Thickness nicht als Kovariable enthalten ist. Grund dafür sind die fehlenden Werte für diese Variable im Kalifornien-Datensatz.

Auch hier fangen wir erneut mit dem Effekt auf den Erwartungswert an. Anhand von Abb. E lässt sich der Effekt auf den Erwartungswert der triggernden Magnitude mit dem der completeness Magnitude vergleichen.



Multiplikativen Effekt auf den Erwartungswert der getriggerten Magnitude

Magnitude

Abb E: Multiplikative Effekte der triggernden sowie der completeness Magnitude auf den Erwartungswert der getriggerten Magnitude für Kalifornien

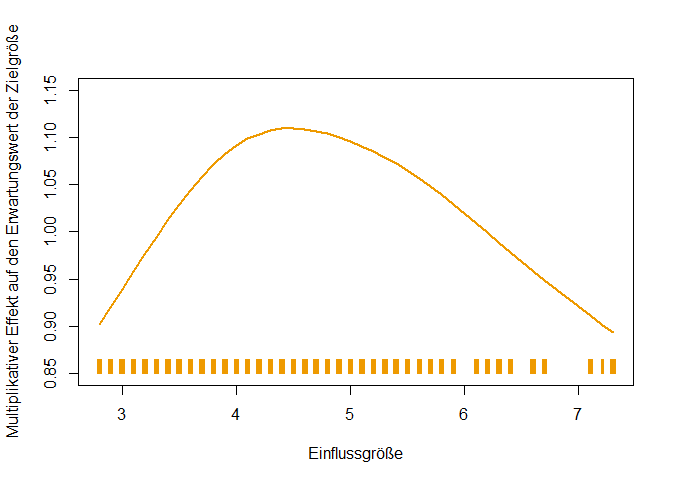
Hier zeichnet sich ein ähnliches Bild wie bereits für Japan ab: die triggernde Magnitude hat zwar einen kontinuierlich ansteigenden positiven Effekt auf die getriggerte Magnitude, dieser ist jedoch im Vergleich zum Effekt der completeness Magnitude äußerst gering. Die Interpretation ist analog zu Japan.

Abb. F: Multiplikativer Effekt der triggernden Magnitude auf die Varianz der getriggerten Magnitude für Kalifornien

Betrachtet man nun den Effekt der triggernden Magnitude auf die Varianz der getriggerten Magnitude (Abb. F), lässt sich bis zu einer Magnitude von ungefähr 4.5 ein steigender Effekt auf den Faktor 1.10 beobachten. Anschließend fällt der Effekt auf die Varianz wieder auf sein Ursprungsniveau von 0.90 zurück. Die triggernden Magnitude hat somit zuerst einen ansteigenden und positiv werdenden Effekt auf die getriggerte Magnitude. Eine hohe triggernden Magnitude bewirkt jedoch eine Abnahme der Varianz der getriggerten Magnitude. Die Stärke der getriggerten Beben findet sich in einem kleineren Bereich wieder.

Man stellt fest, dass die completeness Magnitude einen um ein vielfaches stärkeren Einfluss auf die Varianz der getriggerte Magnitude hat, als die triggernde Magnitude, siehe Abb. G

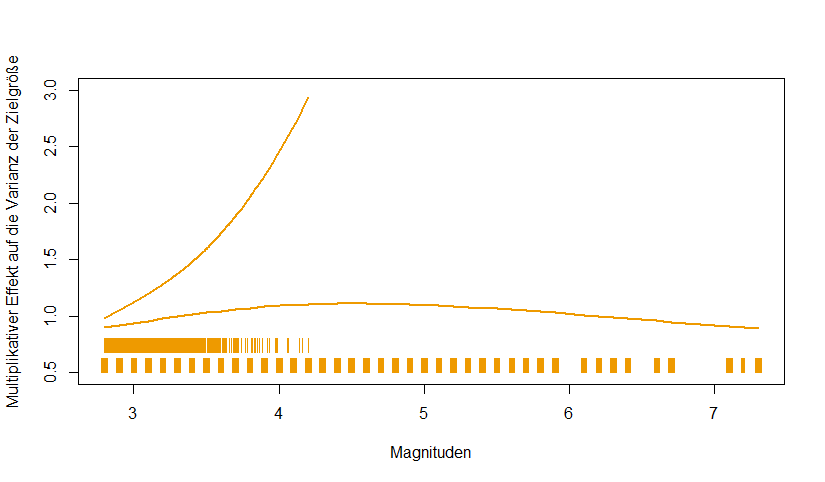


Abb. G: Multiplikative Effekte der triggernden sowie der completeness Magnitude auf die Varianz der getriggerten Magnitude für Kalifornien

Mit steigender completeness Magnitude beobachtet man eine Verdreifachung der Varianz über den gesamten Ausprägungsbereich. Dieses Ergebnis ist in Hinsicht auf die zuvor gegebene Erklärung bezüglich der Trunkierung der Daten unerwartet. Durch eine completeness Magnitude ungleich 2.8 befindet man sich in einer Shortterm-Incompleteness, also eine Phase, in der nur Beben die stärker als die aktuelle completeness Magnitude sind, gemessen werden. Der Bereich der gemessenen Magnitude wird also mit steigender completeness Magnitude geringer, man erwartet somit einer Verkleinerung der Varianz, was hier nicht der Fall ist. Der beobachtete Effekt kann nicht von der completeness Magnitude erklärt werden. Man sollte diesen Effekt mit Vorsicht interpretieren und die Daten auf einen strukturellen Zusammenhang untersuchen, der nicht durch das hiesige Modell erklärt werden kann.

1. Unterscheiden sich die Beobachtungen für Erdbeben-Daten aus Japan und Kalifornien?

1. https://de.wikipedia.org/wiki/Erdbeben [↑](#footnote-ref-1)
2. https://de.wikipedia.org/wiki/Plattentektonik [↑](#footnote-ref-2)
3. https://www.geothermie.de/bibliothek/lexikon-der-geothermie/k/kontinentalplatte.html [↑](#footnote-ref-3)
4. https://de.wikipedia.org/wiki/Plattentektonik [↑](#footnote-ref-4)
5. https://de.wikipedia.org/wiki/Transformstörung [↑](#footnote-ref-5)
6. https://de.wikipedia.org/wiki/Erdbeben [↑](#footnote-ref-6)
7. https://de.wikipedia.org/wiki/Seismograph [↑](#footnote-ref-7)
8. https://de.wikipedia.org/wiki/Magnitude\_(Erdbeben) [↑](#footnote-ref-8)
9. https://de.wikipedia.org/wiki/Richterskala [↑](#footnote-ref-9)
10. https://de.wikipedia.org/wiki/Erdbeben [↑](#footnote-ref-10)
11. https://en.wikipedia.org/wiki/Gutenberg-Richter\_law [↑](#footnote-ref-11)
12. quelle [↑](#footnote-ref-12)
13. https://de.wikipedia.org/wiki/Erdbeben [↑](#footnote-ref-13)
14. Email von Christian Grimm am 17.05.2021 [↑](#footnote-ref-14)
15. Masterarbeit [↑](#footnote-ref-15)
16. Fahrmeir [↑](#footnote-ref-16)
17. Fahrmeir 90 [↑](#footnote-ref-17)
18. Manthei, Gerd: Anwendung des Gutenberg-Richter-Gesetzes in der Schallemissionsanalyse, S. 1 [↑](#footnote-ref-18)
19. Meintrup seite 275 [↑](#footnote-ref-19)
20. Masterarbeit 11 [↑](#footnote-ref-20)
21. Fahrmeir 303 [↑](#footnote-ref-21)
22. Fahrmeir 308 [↑](#footnote-ref-22)
23. Simpson, Gavin: Modelling seasonal data with GAMs, 2014 https://fromthebottomoftheheap.net/2014/05/09/modelling-seasonal-data-with-gam/ [↑](#footnote-ref-23)