

【研究ノート】

ビッグデータを用いた地震データサイエンスの試み

石 塚 隆 男

A Study on Seismic Data Science using Chronological Big Data

ISHIZUKA, Takao

Abstract

This article analyzes interlock properties between faraway earthquakes using chronological catalog of world earthquake occurrences. Earthquakes occurs by plate tectonics, movement of magma, fault slip, and these combined causes. Earthquake energy is massive, and seismic waves propagate to the opposite side of the earth. We assume that consecutive earthquakes are affected by prior earthquakes. In other words, megaquakes add stress to fragile faults in the process of propagation. Fifty-years USGS earthquake catalog data were downloaded, and paired earthquake data, i.e. bigram data were constructed. Bigram is one of the popular methods in natural language processing. It became evident that the occurrence of earthquakes is not random but has some trends by analyzing the frequency of paired seismic center. In seismology, many researchers are engaged in clarification of physical mechanism of past earthquakes. Seismic data science using chronological big data appears promising.

Key Words

USGS, bigram

目 次

- | | |
|----------------|-------|
| 1. はじめに | 4. 結果 |
| 2. 地震間の連動性のモデル | 5. 考察 |
| 3. 研究方法 | |

1. はじめに

わが国では頻繁に地震が発生し、尾池(2007)によれば、1980年代から日本は活動期に入ったとされている。大規模地震は経済や生活、生命に甚大な損失をもたらす。グローバル・サプライチェーンに大きく依存している現代、遠隔地の地震にサプライヤーが被災し、企業の生産活動が停止する事態に追い込まれることも珍しくない。日本はいつでもどこでも地震が起きてもおかしくない中で、企業や個人が労働や居住の場所について自らデータや情報を収集し判断しないと生命や財産を守れなくなりつつある。

わが国では、1960年代から地震学者により東海地震の可能性が指摘され、政府は精密な観測網を整備してきた。しかしながら、2021年に至るまで東海地域を震源とする地震は起きておらず、1995年の阪神・淡路大震災(兵庫県南部地震)、2004年の新潟県中越地震、2011年の東日本大震災(東北地方太平洋沖地震)、2016年の熊本地震、2018年の北海道胆振東部地震に見るように政府の監視体制以外の地域でM6以上の地震が発生している。地震予知について日本地震学会は2017年に「時間を指定した短期的な予知は、現在の技術では困難」との見解を公表している。2021年3月に政府の地震調査研究推進本部は主要活断層帯で発生する地震や海溝型地震の長期的な発生可能性として「今後30年以内に震度6弱以上の地震が起きる確率」を公表している。現状では地震予知は困難である以上、せめていつ来るかわからない地震に備えよというのが政府のメッセージであろう。

日本地震学会の2020年度秋季大会のプログラムによれば、地震発生・活動の物理(63件)、強震動・地震災害(37件)、津波(25件)、活断層・歴史地震(21件)、地震予知・予測(7件)の報告がなされている。また、プレートの境界面のメカニズムの解明など過去の地震の地域を対象とした研究も多く、個別あるいはローカルな地震を対象に余震が起きた原因や断層のズレが生じた原因を探る研究報告がなされている。地震は世界各地で起きており、地球内部のマグマの動きに関係することから地震のエネルギーを地球規模で俯瞰的にとらえる必要性が示唆される。2021年12月3日には山梨県と和歌山県で相次いで震度5弱の地震が観測されたが、地震学者は2つの震源の距離が離れており、時期が重なったのは偶然であり、無関係であると述べている(朝日新聞(2021))。しかし、広域を揺らす地震のエネルギーは爆発事故の比ではなく、地震波が地球内部を伝わり、他の地震に作用すると考えるのが自然であろう。

そこで本稿は、米国地質調査所(USGS)が公開する地震カタログのビッグデータを用い、地震データサイエンスにより巨大地震間の連動性について知見を得ることを目的とする。

今回、地震間の連動性を調べるために、自然言語処理の手法であるbigramの考え方を応用し、分析を行った結果、連動性が示唆される知見が得られたので報告する。本稿は石塚(2020)(2021)をもとに加筆修正を行った。

2. 地震間の連動性のモデル

2.1 先行研究

わが国で起こる地震については、日本列島が四枚のプレートの交差上にあることからプレート説とそれに伴う活断層のズレ、並びに過去の地震の余震等により事後的な説明がなされてきた。近年、プレート境界における脱水脆性化（長谷川(2012)）、水の流れに伴う摩擦力の低下（産総研(2019)）、サーマルプレッシャライゼーション（Hirono(2019)）、熱エネルギーの移動（角田(2016)）などの新説が公表されている。さらに、西村(2011)はプレートについて GPS センサの動きから一体型では説明が困難であるとしてマイクロプレート（ブロック、地塊）の集まりととらえるモデルを提唱している。

単純なプレート説では説明がつかない巨大地震が日本を含む世界各地で発生している現在、データやエビデンスに基づく新たな理論構築と実用化が喫緊の課題となっている。

地震のメカニズムを地質構造ではなく、生起する現象を数理的に扱うモデルについて先行研究を行った。地震の発生は確率事象であることからポアソン過程に代表される点過程による統計モデル（ETAS モデル）が提案され、先行地震による後続地震発生率への影響や異常事象と大地震発生との因果関係の解析が行われている（尾形(2015)）。大きな地震は本震に前震、余震を伴い、時間間隔が不規則であることからセミマルコフ過程をあてはめる研究も行われている。確率過程モデルを実用に供するためには推移確率行列を推定する必要がある。統計的因果推論に基づくグラフィカルモデリングやベイジアンネットワークの活用も考えられるが、地震の発生データを網羅的に収集し、特性を把握することが必要であろう。この他、熱移送説（角田(2016)）や MEGA 地震予測（村井）が知られているが、日本地震学会では ETAS モデル以外の研究報告は見あたらない。

2.2 エンタングルメントモデル

広域を揺らすことができる地震のエネルギーは非常に大きく、遠方の地震波が地殻内を伝わり、他の地震を誘発する可能性は十分ありうる。原(2005)は、2003年の十勝沖地震（M8.1）が 8800km 離れたアリゾナの地震計に明確に記録されたことを紹介している。しかし、世界中で M4.5 以上の地震は 1 年間に 6000 件以上発生しているため、遠方で起きた地震が別の地域の断層と共振し、連動した証拠をつかむことは容易ではない。こうした点を踏まえ、本稿では地震を地球全体の視点からとらえるために「非局所性」の意味で量子力学でいうところのエンタングルメント（もつれ）による解釈を試みる。

地震の発生メカニズムは、法線応力とせん断応力の 2 種類の力のモーメントにより説明できる（纈纈(2020)）。すなわち、断層ができる原因として、地殻が均質でないこと、温度の差により成分の差、流速の差が生ずることが考えられる。断層がずれる原因は、せん断応力 > 法線応力とな

るダブルカップルのインバランスで説明でき、エネルギー保存則により震源のせん断応力のエネルギーが他の断層に伝播する、としている。この2種類のモーメントのバランスが崩れた時にずれ=断層が生ずることから2種類の力のモーメントは、量子の2方向のスピに対応させることができる。地震の発生はストレス=エネルギーを解放し、そのエネルギーは固有振動数が近い他の断層のモーメントとして蓄積される。地震が起こるたびに、他の断層にストレスは蓄積され、2種類の力のモーメントのバランスが崩れると地震が発生すると考えられる。こうして、時間・空間に広がる多重のエンタングルメントネットワークとして地震エネルギーのネットワークが形成されていると想定することにしよう。図1に1970年1月から2021年5月までに世界各地で起きたM4.5以上の地震の震央の分布を示す。プレートの境界で地震が多発していることが読み取れる。図2は地球内部のマントルが対流運動し、上昇流(マンテルプルーム)が溶けた状態であるマグマが地殻やプレートの境界を移動し、火山噴火や断層に圧力を及ぼし、地震を誘発し、さ

図1. M4.5以上の地震の震央の分布(1970.1.1~2021.5.15)

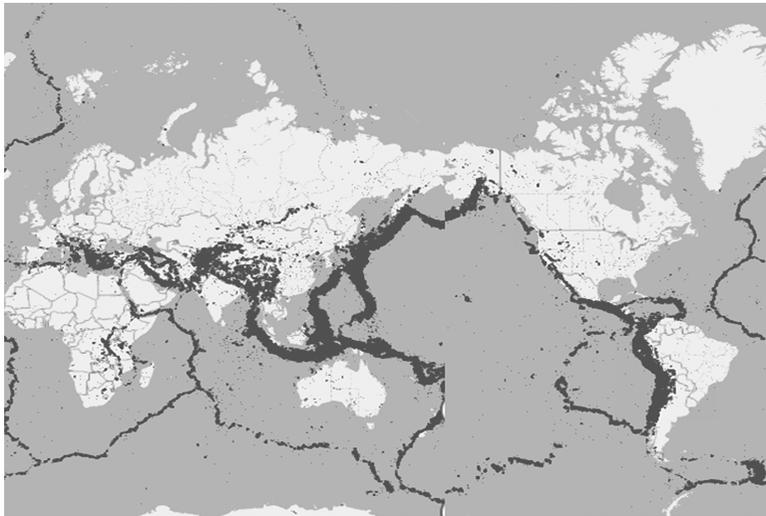


図2. 地震エネルギーのネットワーク

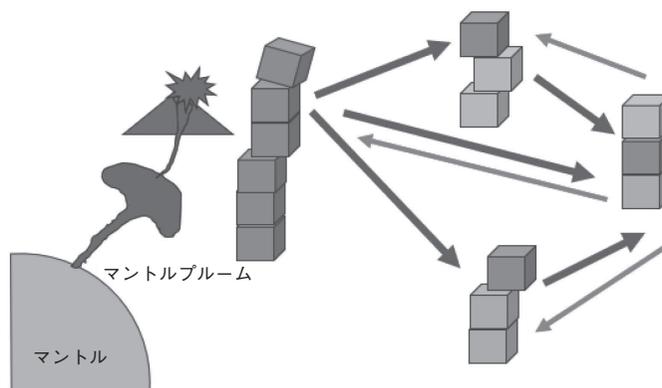


図3. 1970年以降、過去50年間の地震の規模別発生件数の分布

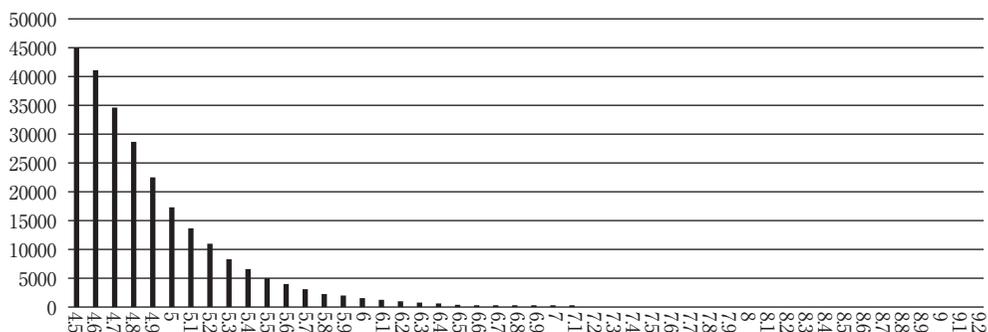
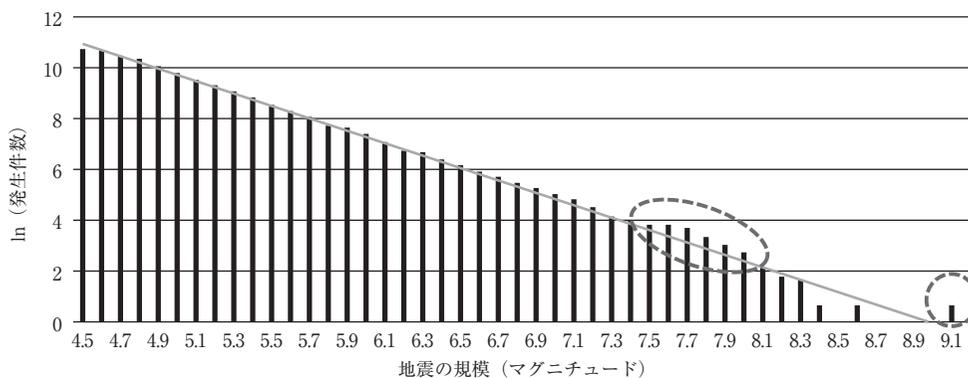


図4. 図3の規模別発生件数の自然対数の分布

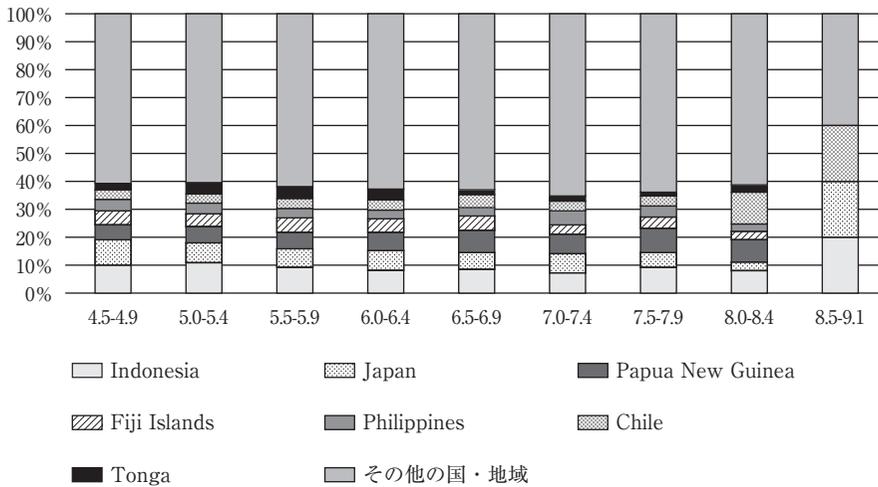


らに他の断層に伝播・蓄積される様を示したものである。

図3に1970年以降、過去50年間の地震の規模別発生件数の分布を、図4に図3の発生件数の自然対数をとった分布を示す。発生件数はほぼ指数関数的に減少していることがわかるが、図4に示すようにM7.7~8.1とM9.1の発生頻度が直線からやや逸脱した大規模災害があることが読み取れる。図5に地震の規模別に国・地域の発生割合を示す。世界でM8.5以上の地震は、日本、インドネシア、チリの3地域で多く、これらの国が太平洋プレートの境界にあり、全体の約60%を占めていることがわかる。

運動性のモデルとして神経回路のように地震を起こす断層が相互につながった断層ネットワークが考えられるが、本稿ではマグニチュード4.5以上の地震が年間数千件発生していることを考慮し、先行する巨大地震と後続地震の間の相関関係に注目することにする。地震のエネルギーは地表面の局所にとどまらず、地球全体の断層に影響（ストレスや歪み）を与えることから、地震間の相関性に関する仮説として、「ひとつの地震には先行する複数の地震が絡み合う」を検討する。局所的な地震現象は、マクロ的に非局所性=エンタングルメントに因ると考えられる。

図5. 地震の規模別, 国・地域の構成割合



3. 研究方法

3.1 対象データ

長期間にわたり, 広範囲に網羅的に地震の発生状況を収集する必要から, USGS (米国地質調査所) の Search Earthquake Catalog を用い, 1970年1月1日から2021年3月5日までの約51年間の全世界で起きたマグニチュード4.5以上の地震250,654件について, 発生日時, 震源の位置(経度・緯度), 深さ, マグニチュード, 震央の場所, 地域(国)をデータベース化した。地域の表現に表記ゆれが見られたため, 集計上, 国名または地域名に統一するようにした。

3.2 bigram 分析

地震事象は不等間隔時系列であり, 2つの地域で起きた地震は対応のあるデータではないため, 通常の相関を計算することはできない。そこで, 任意の地震発生後, 一定の期間内に世界中で発生したM6.0以上のすべての地震との対をつくり, 分析を行うことにした。

地震間の関連性や連動性を分析するため, 自然言語処理で用いられる bigram の考え方を活用する。対の頻度を調べることによりある地域で地震が起きると一定の期間内に特定の地域で地震が起きる確率がどの程度かを計算することができる。N-gram とは, 連続するN個の文字, 単語の塊を指し, N=2の場合, bigram とよぶ。地震の bigram の構成方法として,

方法1: ある地震と次に起きた地震で対(ペア)をつくる

方法2: ある地震から○日以内に起きた地震で対をつくる

が考えられる。方法1は, マルコフ過程であるが, 震源間の距離をどう扱うか問題がある。方法2は, 方法1の問題に加えて, ひとつの地震が複数回カウントされる問題がある。

本稿ではランダムで小規模の地震ではなく, 繰り返し起こる固有地震の間の相関やもつれを分

図 6. bigram のクロス集計イメージ

後事象の地域→ ↓前事象の地域	日本	バヌアツ	…	j	…	チリ	TOTAL
日本							
バヌアツ							
…							
i				X _{ij}			X _{i.}
…							
チリ							
TOTAL				X _{.j}			X _{..}

析するために M6.0 以上の地震を対象とした。

地震の本震と余震の時間間隔は地震学者の見解によると考えられ、40 年前の地震の余震と発表されることもある。本稿では、ある地震から 60 日の期間内に起きた M4.5 以上の地震で対を構成し、bigram データとした。60 日以内の地震を抽出することは、先行地震を input 系列とし、ウィンドウ幅が 60 日の長方形ウィンドウのフィルタを通したことに相当する。

bigram データは、前事象の地域と後事象の地域のクロス表の形で起きた地震の数を集計することができる (図 6)。実数値のままでは値の大小を評価できないため、タテ計の % とヨコ計の % の積により評価することにした。

地域 i で起きた地震の後 60 日以内に、地域 j で起きた地震の件数を X_{ij} とする。地域 i で 60 日以内に複数回の地震があると X_{ij} には第 j 地域のひとつの地震が複数回カウントされることになるが、後事象の第 j 地域の総件数 $X_{.j}$ も増えることになる。前事象の第 i 地域の全期間 (約 51 年間) の M4.5 以上の地震の合計を $X_{i.}$ とし、 i, j のすべての組合せについて、

$$\sqrt{\frac{X_{ij}X_{ij}}{X_{i.}X_{.j}}} = \frac{X_{ij}}{\sqrt{X_{i.}X_{.j}}} \quad (1)$$

の値を計算した。

式(1)は、ベクトル i と j のコサイン類似度に相当することから地域 i と j の相関係数に相当する指標である。なお、 $X_{ij}=1$ のセルが多いため、通常の特化係数は使えないが、前事象と後事象が独立ならば、地域数を n とすれば、 X_{ij} の期待値は $1/n \times 1/n = 1/n^2$ であるから、式(1)と $1/n$ を比較することにより特化の程度を評価することができる。

4. 結果

4.1 連動性指標による全体評価

bigram を地域によりクロス集計し、式(1)の連動性指標を算出した結果の一部を表 1 に示す。日本で起きた地震の先行地震の震源地として、インドネシア、パプアニューギニア、チリの割合が高いことがわかる。いずれの地域も日本と同様に地震多発の地域であるため、偽相関の可能性もある。

表 2. M8 地震間の M7 以上の bigram の頻度

bigram の組合せ	頻度
Indonesia ⇒ Japan	5
Philippines ⇒ Japan	5
Fiji Islands ⇒ Indonesia	4
Indonesia ⇒ Vanuatu	4
Papua New Guinea ⇒ Alaska	4
Philippines ⇒ Papua New Guinea	4
Solomon Islands ⇒ Indonesia	4
Chile ⇒ Papua New Guinea	3

表 3. 頻度 2 以上の trigram

trigram の組合せ	頻度
Kuril Islands ⇒ Philippines ⇒ Japan	2
Mexico ⇒ Papua New Guinea ⇒ Alaska	2
South Island of New Zealand ⇒ Indonesia ⇒ Japan	2
Aegean Sea ⇒ Kermadec Islands ⇒ Philippines	1

震災の間に起きた M7 クラスの地震の震央を発生日時順に矢印で結んだものである。3.11 のほぼ 1 年前にチリで巨大地震があり、ニュージーランドやインドネシア付近でいくつかの M7 クラスを経た後、3.11 が起きたことがわかる。2011 年 2 月 22 日にニュージーランドのクライストチャーチで起きた地震 (M6.1) の 17 日後に 3.11 が起きたため、当時、ニュージーランドで地震が起きると日本で地震が起きるのではないかという噂が広まったが、図 7 から事実それほど単純ではないことがわかる。いずれの巨大地震も太平洋プレートの境界面で起きており、エンタングルメントが疑われるが因果関係があるかどうかは明らかではない。

表 2 よりインドネシアとフィリピンの地震に続き、日本で M7 以上の地震が起きる頻度が高いことがわかる。ひと月に M7 が 4 回あると 1 ~ 2 か月後に M8 クラスが起きている。bigram だけでなく、3 つの地震の連関性を調べるために trigram についても集計した結果を表 3 に示す。日本が 3 つ目に来るケースはクリル島⇒フィリピン⇒日本、並びに、ニュージーランド南島⇒インドネシア⇒日本の 2 系統があることがわかったが、bigram ほどの頻度には至らず、地震の発生メカニズムが複雑であることを伺わせる。

4.3 先行地震の震源域別の後続地震の分布状況

先行地震の震源域により日本で起きる後続の地震の震源域の分布が異なるかどうかを bigram データから調べ、散布図を作成した。図 8 にパプアニューギニアが先行地震の震源域の場合の散布図を示す。

図 8 より、フィリピン海プレートに沿って地震が多発しており、南海地震へのトリガーとなる可能性が示唆される。フィリピン、インドネシア、ニュージーランド、チリを震源域とする先行

図8. パプアニューギニアを震源域とする M7 以上の地震の 60 日以内に日本で起きた M6 以上の地震

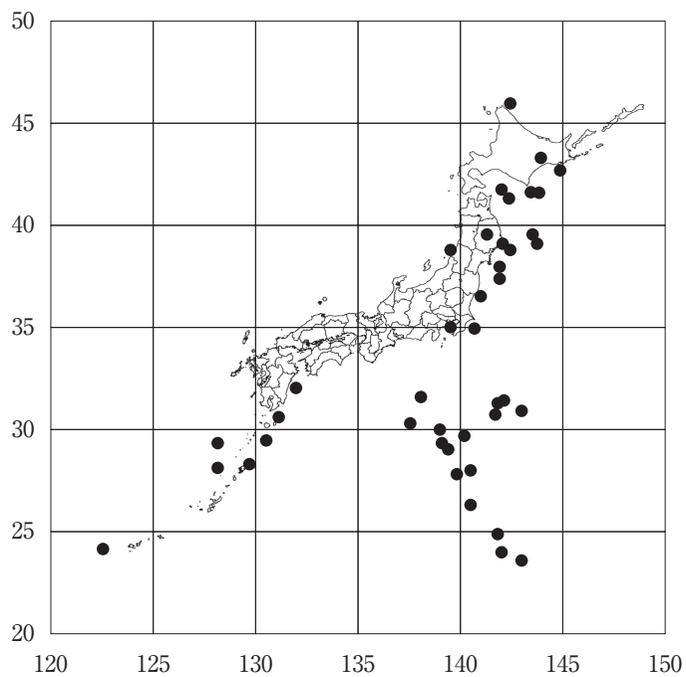
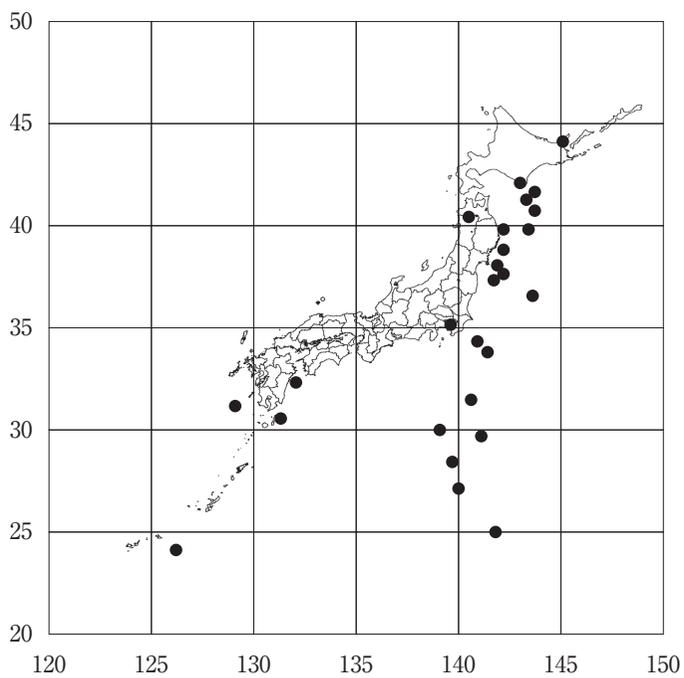


図9. チリが震源の M7 以上の地震の 60 日以内に日本で起きた M6 以上の地震の震央の分布



地震について同様の散布図を作成した結果、先行地震の震源域により日本で起きる後続地震の分布が異なることが目視で確認された。

図9にチリが震源のM7以上の地震の60日以内に日本で起きたM6以上の地震の震央の分布を示す。大半の震央が伊豆諸島から東北に至る太平洋の東側に分布していることがわかる。フィリピンやインドネシアを震源とする地震の60日以内に日本で起きた地震の震央の分布図を描くと、必ずしも太平洋の東側ではなく、九州から東北に至る太平洋の南側に分布している。

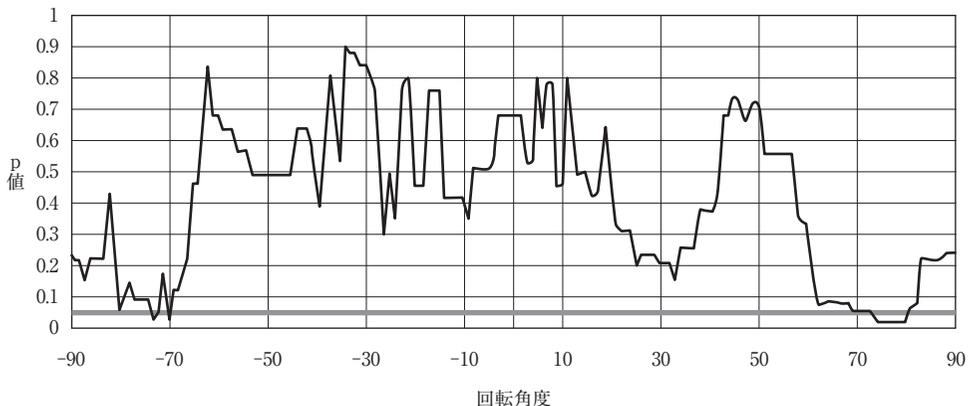
4.4 先行地震による散布図の違いの定量化

遠隔地の地震が日本で起きる地震と連動しているならば、後続地震の散布図は先行地震の震源域で異なることが予想される。逆に、先行地震の震源域と無関係に日本で地震が起きるならば、散布図は同等とみなせるだろう。実際に相関係数や2次元平面の領域にプロットされた点の数の適合度検定や比率の差の検定では有意にならなかった。

本研究では2つの散布図間の類似性を定量化するために、震央の経度と緯度を標準化・回転し、周辺分布として新たな経度の標準化値を $-1, 0, +1$ により4区分し、カテゴリーデータ化した。回転角度を変化させることにより、適合度検定に用いるカイ2乗統計量からp値を求め、パプアニューギニア vs. チリ、フィリピン、ニュージーランド、インドネシアの散布図間でp値が0.05未満になることが確認された。図10に先行地震の震源域がパプアニューギニアとチリの散布図間の違いについて回転角度が $-90^\circ \sim +90^\circ$ のp値の変化を示す。

図10より回転角度が 74° 付近でカイ2乗統計量が有意となることがわかる。少なくともパプアニューギニアを先行震源域とする地震は今回の比較対象地域の地震とは後続地震の分布が異なることが示された。

図10. 散布図を回転し、周辺分布のカテゴリー化によるカイ2乗統計量のp値
(パプアニューギニア vs. チリ)



4.5 地震の静穏期の可視化

地震は忘れた頃にやって来ると言われるように、しばらく穏やかな日々が続いた後に突然、大地震が起きることがある。そこで、地震の規模と地震の前日に M4.5 以上の地震が起きた地域数の散布図を作成した (図 11)。図 11 に破線により包絡線を書き加えると地震の規模が M6.5 を超

図 11. 地震の規模と前日の M4.5 以上の地震の地域数との関連

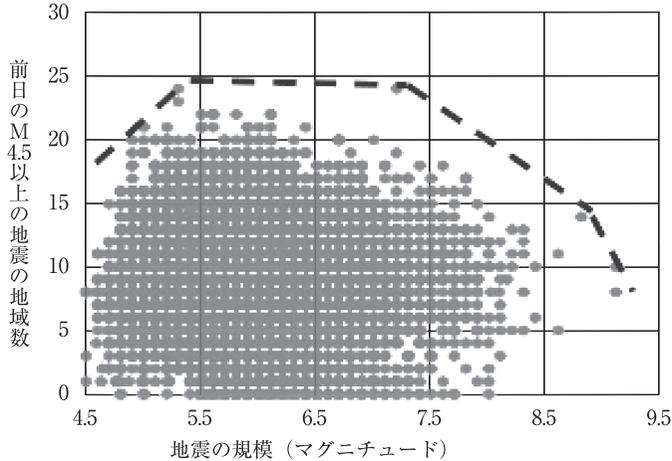


図 12. 1970 年以降の熊本地震 (経度 131°, 緯度 32°) の発生年月日と規模との関係

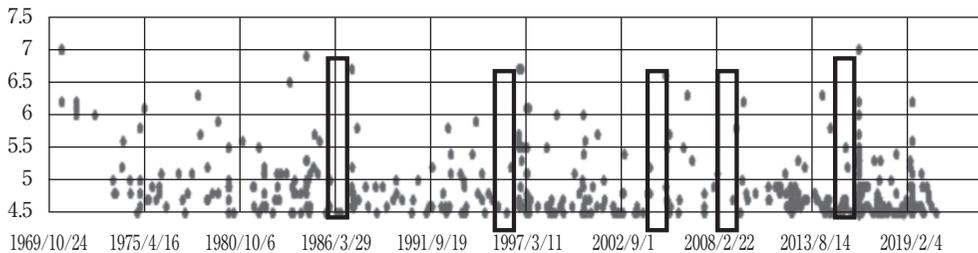
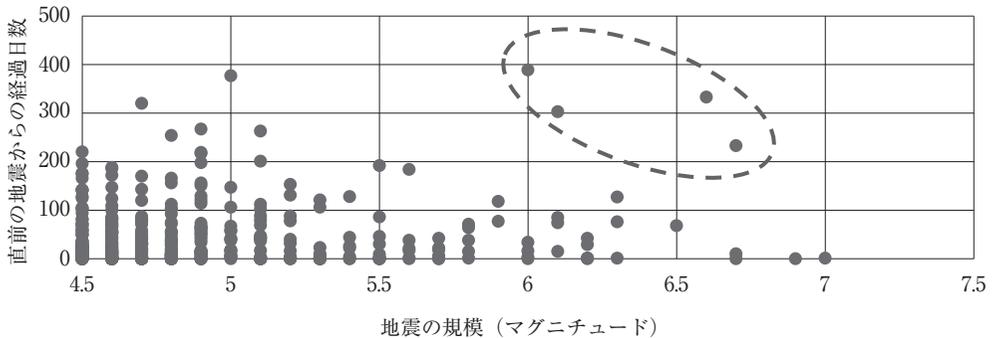


図 13. 1970 年以降の熊本地震 (経度 131°, 緯度 32°) の地震の規模と直前の地震 (M4.5 以上) からの経過日数との関係



えたあたりから前日に起きた地震の地域数はやや小さくなる傾向にあることがわかる。大きなカタストロフィーが起きる際にはその部分に力が集中する結果、静穏な地域が増えることが伺える。

図 12 に 1970 年以降の熊本県を震源とする地震について横軸に年月日、縦軸に地震の規模をとり、プロットした結果を示す。M6.5～7.0 クラスの地震の前に静穏期があることがわかる。図 13 は 1970 年以降の熊本県を震源とする地震の規模と直前の地震（M4.5 以上）からの経過日数との関係を示したものである。静穏期がしばらく続き、巨大地震が起きていることがわかる。

5. 考察

遠隔地の地震により別の地震が引き起こされるかどうかについては地震年表データを分析した限りでは因果関係があるのか偶然の一致なのかは明らかにはならなかった。たとえば、日本、インドネシア、チリは地震が多発する国であり、今回の bigram では 60 日以内に 1 回の M6 があればカウントするので集計の単位を日本全体ではなく、細分化することも必要であろう。大きな余震の前の余震回数が大森公式よりも減少することは松浦(1986)により明らかにされている。比較的長い静穏期こそ注意を喚起すべきであろう。

地震の原因については、プレートテクトニクス説が知られているが、プレートの移動を起こすマントル対流について不明な点も多い（岩森(2016)）。角田(2016)は、マントルから噴き上がったマグマの熱の通り道に沿って地震が起きるとする熱移送説を提唱している。先行地震と後続地震の相関関係には隠れた変数が作用している可能性も考えられる。

6. 結語

地球規模で、地震イベントのデータを収集し、対を作成し、地域間の関連の強さを可視化した結果、日本に関しては、先行地震の地域により日本列島のどこで地震が発生するかに関して異なることが明らかになった。現在の科学技術では地震予知は困難であるが、先行する地震の発生により近いうちに地震が起きる可能性を検知できれば、多少の備えができると思われる。

しかし、USGS の地震カタログデータが全世界をカバーしたのは 1970 年以降であり、50 年分のデータで大規模地震の発生件数は限られるため、サンプル数は十分とはいえない。

今回、bigram を構成するためのウィンドウの幅を 60 日としたが、本震と余震の時間間隔は数年の場合もある。しかし、ウィンドウの幅をいたずらに大きくすることは無関係な地震まで多く拾うことになりかねない。bigram は単純なモデルであり、各地の地震の履歴情報は活用されていない。地震の連動性を分析するためにより精緻なモデル化が必要である。人間が把握できているのは地表に見えているごく一部の活断層のみであり、潜在的な活断層が無数に存在する。過去の地震の震源の深さをデータとしてどう活かすかも検討する必要がある。また、地震との関わり

において、特定の火山の噴火にのみ注目するのではなく、地球規模で地下のマグマの動きを推測する必要があるだろう。

将来的には、実際の地震データによるデータ同化 (data assimilation) が可能な地球内部～地表の地球シミュレータを構成できれば、地震データサイエンスは大きく進展することになるだろう。

参考文献

- 朝日新聞 (2021) 「相次ぐ地震、メカニズムは 富士五湖・紀伊水道・トカラ列島 過去にも頻発「時期が重なったのは偶然」」2021年12月7日朝刊, p.34
- 石塚隆男 (2020) 「エンタングルメントの発想に基づく地震年表データの分析」『2020年度日本分類学会シンポジウム講演予稿集』, pp.33-36
- 石塚隆男 (2021) 「bigram による巨大地震間の連動性の分析」『2021日本経営工学会春季大会予稿集』, pp.326-327
- 岩森光 (2016) 「マントル対流と全地球ダイナミクス」『火山』第61巻第1号, pp.1-22
- 尾池和夫 (2007) 『新版 活動期に入った地震列島』岩波書店
- 尾形良彦 (2015) 「地震の確率予測の研究—その展望—」『統計数理』第63巻第1号, pp.3-27
- 角田史雄 (2016) 『次の「震度7」はどこか! 熊本地震の真相は「熱移送」』PHP研究所
- 瀬戸一 (2020) 『地震 どのように起きるのか』丸善出版
- 産総研 (2019) 「日本列島直下に沈み込むプレート内の水の挙動がスロー地震発生に関係」https://www.aist.go.jp/aist_j/press_release/pr2019/pr20190130/pr20190130.html
- 西村卓也 (2011) 「測地学的データに基づくプレート間相互作用の解明」『測地学会誌』第57巻, 第1号, pp.1-14
- FAQ 2-3. 地震予知の信頼性 (2017年12月修正) | 公益社団法人 日本地震学会 https://www.zisin.jp/faq/faq02_03.html, 2021年12月12日確認
- 長谷川昭 (2012) 「沈み込み帯の地震の発生機構—地殻流体に規定されて発生する沈み込み帯の地震—」地学雑誌 *Journal of Geography*. 121 (1) pp.128-160
- Hirono, Tetsuro ; Tsuda, Kenichi & Kaneko, Shunya (2019). Role of Weak Materials in Earthquake Rupture Dynamics. *Scientific Reports*, 9
- 原辰彦 (2005) 「地球の地震学的構造」『地学雑誌』 *Journal of Geography*. 114 (3) pp.323-337
- 松浦律子 (1986). Precursory Quiescence and Recovery of Aftershock Activities before Some Large Aftershocks. 東京大学地震研究所彙報 61 (1) pp.1-65
- 村井俊治・地震科学探査機構 (JESEA) 『MEGA地震予測』 <https://www.jesea.co.jp/>