階層ベイズモデルを用いた土砂災害発生数予測モデルの検討

A HIERARCHICAL BAYESIAN MODEL FOR PREDICTING THE NUMBER OF SEDIMENT-RELATED DISASTERS

三好 夢珠[1]・笠間 清伸[2]・李 洋[3]・荘 恵民[4]
Yumemi MIYOSHI, Kiyonobu KASAMA, Yang LI and Huimin ZHUANG

[1]九州大学大学院工学府（〒819-0395 福岡県福岡市西区元岡 744 番地ウエスト 2 号館 11 階 1110 号室）
E-mail: miyoshi.yumemi.129@s.kyushu-u.ac.jp
[2]九州大学大学院工学研究院（〒819-0395 福岡県福岡市西区元岡 744 番地ウエスト 2 号館 11 階 1128 号室）
E-mail: kasama@civil.kyushu-u.ac.jp
[3]九州大学大学院工学研究院（〒819-0395 福岡県福岡市西区元岡 744 番地ウエスト 2 号館 11 階 1111 号室）
E-mail: li.yang.168@m.kyushu-u.ac.jp
[4]九州大学大学院工学研究院（〒819-0395 福岡県福岡市西区元岡 744 番地ウエスト 2 号館 11 階 1110 号室）
E-mail: zhuang.huimin.618@m.kyushu-u.ac.jp

Key Words: Sediment-related disaster*, Hierarchical Bayesian model, precipitation, risk assessment*

1. 背景

近年，気候変動に伴う集中豪雨の頻発・激甚化により，土砂災害による被害リスクが増大している．このような状況を背景に，土砂災害の発生傾向の分析や発生を予測した研究が行われてきた[1)]．しかし，これらの研究の多くは詳細な降水データや地形情報を必要としており，必要なデータの収集や整備に多大な時間と労力を要する．

他方で，土砂災害の発生メカニズムは，降雨に加え，地形・地質といった地域固有の素因が複雑に関係するため，その発生傾向には地域差が大きいことが知られている．

階層ベイズモデルは，複数の階層構造をもつデータに対して，ベイズ推定により事後分布のパラメータを求める手法であり，グループ間の情報を共有しながら，個別の推定精度を高められる特徴を持つ．土砂災害発生数に関する観測データは 41 年分と限られているが，実質的に 41 年×47 都道府県＝1,927 件のデータを活用可能となる．特に，地域特性のような該当するデータのみを局所的に説明する変数（変量効果）を適切に捉える点で有効とされている．他分野においては，猪俣によって林業災害の発生における地域差が，廣井によって地震火災の出火件数に関する地域差が明らかにされた[2,3)]．このように，従来は定量化が困難であった個体差や地域差の評価に活用されている．

そこで本研究では，階層ベイズモデルを用いて都道府県単位の情報から土砂災害の発生傾向を把握・評価するモデルの構築を試みた．全国を対象に都道府県単位で地域差を定量的に評価するとともに，防災施策を実施する行政単位である都道府県ごとの発生傾向を明らかにすることで，効率的な防災対策への活用に寄与することが期待される．

2. 手法

本研究では 1983～2023 年までの各都道府県の土砂災害の発生数[4)]，山地面積[5)]，最大 1 時間降水量

[6]，最大 24 時間降水量[6]を収集した．本研究で用いたデータは，各年度において 47 都道府県すべてを対象としており，41 年間を通じてデータ数の欠落や変動はない．土砂災害は一般に土石流・地すべり・がけ崩れの 3 種類に分類され，本研究ではこれらを区別してモデルに入力した．ここで，地震や火山活動が要因となる災害発生数は除外し，降雨に起因する災害発生数に限定した．

土砂災害の発生に影響を与える因子として，本研究では降水量に着目し，降水の影響を表す変数として，各都道府県の単位面積あたりの年次別の最大 1 時間降水量rc_{ij}と最大 24 時間降水量rd_{ij}を使用した．用いた降水量データは，都道府県ごとに 1 地点を対象とし，各都道府県に設置されている管区気象台，地方気象台および沖縄気象台の観測データを用いた．各都道府県で複数の対象観測所が存在する場合には，都道府県を代表する地点として県庁所在地に設置された観測所のデータを採用した．収集したデータは，入力する説明変数は，最小値と最大値がそれぞれ-1 と 1 となるように標準化を行った．

本研究では階層ベイズモデルを用いた．階層ベイズモデルとは，ベイズ推定を拡張したモデルであり，パラメータθに対して事前分布$p(\theta)$を設定し，データyを観測した後にベイズの定理式(1)を使って事後分布$p(\theta|y)$を推定する．ここで，$P(y|\theta)$は尤度関数，$p(y)$は周辺尤度である．$p(y)$は定数であるため，$p(\theta|y)$に比例すると考えられる式(2)．

$$p(\theta|y) = \frac{P(y|\theta)p(\theta)}{p(y)} \tag{1}$$

$$p(\theta|y) \propto P(y|\theta)p(\theta) \tag{2}$$

階層ベイズモデルでは，さらに事前分布の形を定めるパラメータθに対しても確率分布（超事前分布）を設定する．このパラメータの階層化により，地域差を表す変量効果を推定することが可能となり，他の都道府県の情報も統合して各都道府県の推定を行うことができる．

本モデルでは，災害発生数の確率分布y_{ij}に，負の二項分布（NB 分布）を採用した．土砂災害発生数は非負の整数値であり，発生しない年が多い一方で，豪雨時には多数の災害が発生するため，平均に対して分散が大きい過分散の傾向を示す．負の二項分布は，平均と分散が等しいという制約をもつポアソン分布とは異なり，過分散を表現できるため，このようなデータに適している．モデルの構造は式(3)～(7)の通りである．

表-1 使用した変数・パラメータ

記号	説明
y_{ij}	i 年次における都道府県 j の各災害発生数
s_j	都道府県 j の山地面積[km^2]
t	年次の増加
e^{a_j}	年次による各災害発生数の変化
e^{b_j}	各災害の平均的な発生しやすさ
e^{c_j}	最大 1 時間降水量による各災害発生数の変化
e^{d_j}	最大 24 時間降水量による各災害発生数の変化
α_j	過分散パラメータ
rc_{ij}	i 年次の都道府県 j の年間最大 1 時間降水量 [mm]
rd_{ij}	i 年次の都道府県 j の年間最大 24 時間降水量 [mm]

$$y_{ij} \sim NegativeBinomial(\mu_{ij}, \alpha_j) \tag{3}$$

$$\mu_{ij} = s_j * exp(\eta_{ij}) \tag{4}$$

$$\eta_{ij} = a_j t + b_j + c_j * rc_{ij} + d_j * rd_{ij} \tag{5}$$

$$a_j, b_j, c_j, d_j \sim Normal(\mu, \sigma) \tag{6}$$

$$\log(\alpha_j) \sim Normal(\mu_{log\alpha}, \sigma_{log\alpha}) \tag{7}$$

ここで，y_{ij}は i 年次における都道府県 j の各災害発生数を表す．y_{ij}は平均μ_{ij}，過分散パラメータα_jをもつ負の二項分布に従うと仮定した．μ_{ij}は都道府県の面積s_jをオフセット項として考慮し，対数リンク関数を用いて線形予測子η_{ij}から求めた．線形予測子は，年次変化を表す係数a_jと年次の変数t，切片b_j，時間降水量係数c_j，24 時間降水量係数d_jにより構成した．また，a_j, b_j, c_j, d_jは，都道府県ごとのばらつきを考慮するため，それぞれ共通の平均μおよび標準偏差σをもつ正規分布に従う階層構造を仮定した．過分散パラメータα_jについても都道府県ごとのばらつきを考慮し，対数スケールで正規分布に従う階層事前分布を設定した．ここで，$\mu_{log\alpha}$および$\sigma_{log\alpha}$は，それぞれ$\log(\alpha_j)$の平均および標準偏差を表すハイパーパラメータである．

1983～2018 年（以降，学習期間とよぶ）のデータを用いてパラメータを推定し，2019～2023 年（以降，検証期間とよぶ）における土砂災害発生数の予測を行いモデルの適合度の検証を行った．モデルの予測には，ベイズ推論によって得られた各パラメータの事後分布を使用し，式(3)～(5)用いて負の二項分布を用いた乱数サンプリングにより予測分布を生成した．各変数，パラメータの説明は**表 1** の通りである．パ

ラメータe^{a_j}は年次による土砂災害発生数の変化を表しており，気候変動による降水パターンの変化，土地利用の変化，治山・砂防事業の進展，観測・報告体制の変化など複数の要因が反映される可能性がある．パラメータe^{b_j}は，都道府県の平均的な災害の発生しやすさであり，年次の経過や降水量に影響されない地形や地質などその都道府県の持つ潜在的な災害発生ポテンシャルを表している．

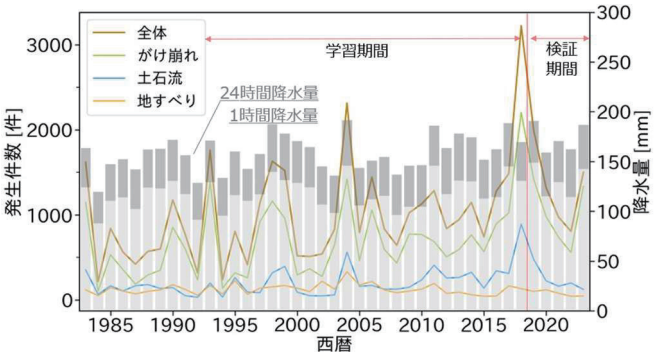


図-1 がけ崩れ，土石流および地すべりの発生傾向

3. 結果

(1) がけ崩れ，土石流および地すべりの発生傾向

図-1 に全国における 1983～2023 年のがけ崩れ，土石流および地すべりの発生数と最大 1 時間降水量，24 時間降水量の都道府県平均値の推移を示す．都道府県の年平均発生数は，がけ崩れで 15.10 件，土石流で 4.47 件，地すべりで 2.56 件であった．土砂災害のうち最も発生数が多いのはがけ崩れであり，1993 年や 2004 年のように降水量の増減に伴って，発生件数が大きく変動する傾向が見られた．土石流の発生件数はがけ崩れに比べて多くはないが，2004 年や 2018 年にみられるようにがけ崩れとともに突発的な増加があった．地すべりは最も発生数が少なく，降水量の増加に伴う急激な発生数の増加も少なかったため，ほかの災害と異なる傾向にあることが示唆された．

長期的な年次の経過による発生数の変化に着目すると，がけ崩れと土石流では年次の経過に伴って増加傾向にある．一方，地すべりは年次の経過による発生数の変化は見られなかった．この傾向の違いはモデルにおけるパラメータe^{a_j}に影響を与える可能性がある．

また，学習期間の最終年にあたる 2018 年では全国で 3000 件を超える土砂災害が確認された．図-1 では各都道府県の降水量平均値を表示しているため，降水量の増加が顕著でないものの，実際に多く災害が発生した西日本では記録的な長期豪雨が確認された．そのため，2018 年の土砂災害発生には降水量の影響，特に最大 24 時間降水量e^{d_j}の影響が大きく学習される可能性がある．

(2) モデルの検証

モデルの適合度を検証するため，検証期間における全都道府県の実測値と予測分布についてカイ二乗適合度検定を行った．図-2 に各災害の検証期間における実測値(N=235)と予測値(N=235,000)のヒストグラム

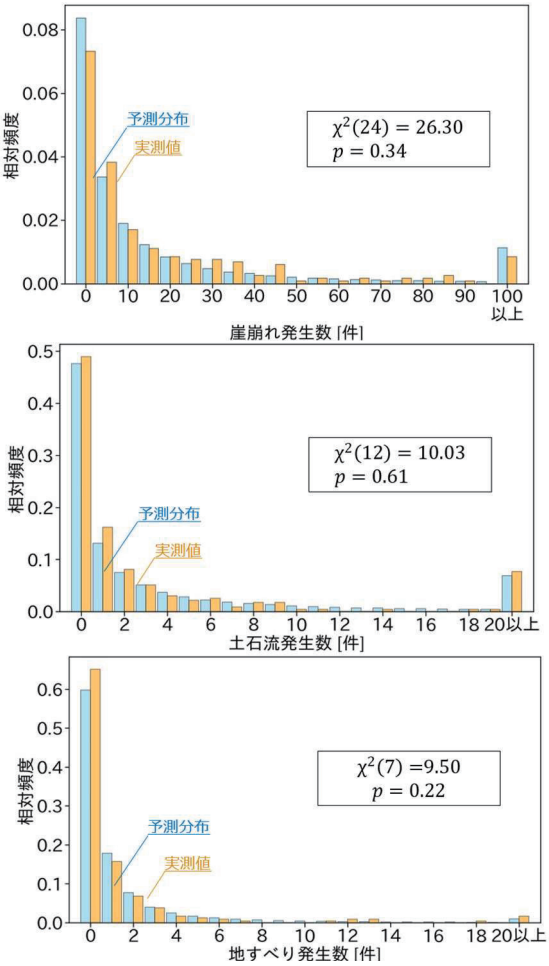


図-2 検証期間における実測値と予測分布（上：がけ崩れ，中：土石流，下：地すべり）

を示す．予測値は，各年・各都道府県について事後分布から 1000 件の予測サンプルを生成した．この予測分布から，各発生件数の相対頻度（発生確率）を算出し，その相対頻度に実測値のデータ数を乗じることでカイ二乗検定に用いる期待度数を求めた．カイ二乗適合度検定の結果，がけ崩れで$\chi^2(24) = 26.30,\ p = 0.34$，土石流で$\chi^2(12) = 10.03,\ p = 0.61$，地すべりで$\chi^2(7) = 9.50,\ p = 0.22$となった．すべて

の災害種別において$p > 0.05$であったことから，予測モデルは 95%信頼水準において実測値と有意な差は認められず，実測値の分布を適切に再現していることが確認された ．

(3) パラメータの推定結果

推定したパラメータe^{a_j}の分布を**図-3** に示す．経年的に災害発生数が増加傾向（事後平均が正の値）を示す都道府県を赤，減少傾向（事後平均が負の値）を示す都道府県を青で表した．また，95%信用区間に 0 を含まない都道府県についてハッチングを行った．つまり，ハッチングを行った都道府県では統計的に有意に土砂災害の発生数に増加傾向/減少傾向があることを意味する．

がけ崩れにおいては，平均値で見ると栃木県，岡山県，島根県，佐賀県を除く 43 都道府県で増加傾向であったが，実際に統計的に有意と言えるのは 21 県であった．土石流においても同様に，平均値では宮崎県，佐賀県，山形県，愛知県，島根県を除く 42 都道府県で増加傾向だったが，統計的に有意と言えるのは 11 県であった．がけ崩れと土石流の両方で有意に増加傾向となったのは，東京都，新潟県，富山県，静岡県，鳥取県，広島県，愛媛県，福岡県の 8 つであった．一方で地すべりでは，平均値では新潟県，富山県，山形県，鹿児島県，福岡県，愛媛県，福井県，鳥取県を除く都道府県で減少傾向であり，11 県で統計的に有意に減少するという結果が得られた．

例外的に，福島県と千葉県の経年的な土石流発生増加数，青森県の経年的な地すべり発生増加数などは，平均値は大きいものの信頼区間も大きいため有意差なしという結果となった．

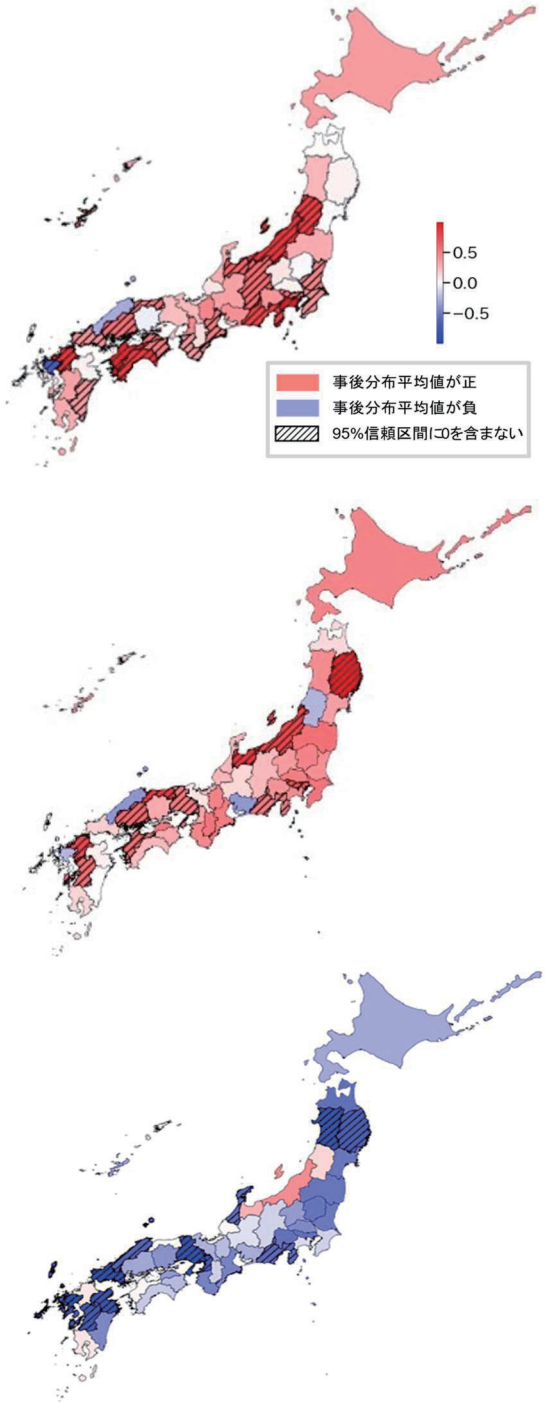


図-3 パラメータe^{a_j}の分布（上：がけ崩れ，中：土石流，下：地すべり）

4. 福岡県と宮城県における詳細分析

(1) モデルの評価

すべての災害種類においてカイ二乗統計量は$p > 0.05$であったことから，本モデルが，各土砂災害の全国的な発生頻度を統計的に十分に再現できたことが示された．本章では，モデルが良好に適合した例として福岡県を，十分に適合していなかった例として宮城県における詳細な分析を行った．

検証期間における福岡県と宮城県の土砂災害発生数の実測値と予測分布を**図-4** に示した．福岡県では，検証期間においていずれの災害も実測値が予測分布内に収まることが確認された．

宮城県では，2019 年の土石流を除き，いずれの災害種別においても実測値が予測分布内に収まった．2019 年には，最大 1 時間降水量 63.5mm，最大 24 時間降水量 303.5mm の記録的な降雨が観測されており，モデルにおいても降雨量の増加に伴い土石流発生数の推定値が増加した．しかし，土石流発生数の最大推定値が 47 件であったのに対し，実際には 146 件の土石流が発生しており，大きく過小評価する結果となった．これは，宮城県では学習期間中に同規

模の土石流災害事例が存在しないため，このような極端な発生数に対する十分なパラメータ推定ができていなかったことが要因であると考えられる．また，土石流と同様に発生数の多いがけ崩れでは，2019 年および 2022 年の発生数の急激な増加については，予測分布の中央値からは外れたものの，予測分布の上限内部には収まっていた．このことから，土石流においても今後データの収集が進むにつれ精度の向上が期待される．

図-5 に宮城県における降水量パラメータの推定結果を示す．**図-5** に示されたパラメータの事後分布を見ると，災害種別によって降水量の影響度に違いが認められた．特に，土石流では降水量に関するパラメータe^{c_j}およびe^{d_j}の平均値が比較的大きく，1 時間降水量および 24 時間降水量が土石流の発生件数に強く影響することが示唆された．一方，地すべりではこれらのパラメータの平均値が相対的に小さく，他の災害種別と比較して降水量の影響が小さいことが示唆された．また，**図-4** に示した宮城県における年次別の実測値および予測分布を見ると，降水量が比較的多かった 2019 年および 2022 年には，土石流の予測値が増加していることが確認された．これに対し，地すべりでは 2019 年および 2022 年における発生件数および予測値の増加は限定的であった．以上のことから，**図-5** で示された降水量関連パラメータの違いが，**図-4** に示される各災害種別の年次別発生件数および推定件数の違いとして反映されていることが示唆された．また，土石流では降水量関連パラメータの分散も大きく，降水量の影響の推定には他の災害種別よりも大きな不確実性が伴うことが示された．

さらに，特定の気象条件におけるモデルの応答性を詳細に評価するため，2023 年の最大 24 時間降水量：97.0mm を基準とし，同年の 24 時間降水量を一律で+20.0mm 変化させた場合の試算を行った．土石流の発生数は，実際の降水量を用いた予測期待値が 1.5 件(95%予測区間上限:10 件)であったのに対し，24 時間降水量を 20mm 増加させた試算では，予測期待値が 2.0 件（95%予測区間上限：13 件）へと増加した．これに対し，24 時間降水量の影響が最も小さかった地すべりでは，予測期待値が 0.7 件（95%予測区間上限：4 件）から 0.8 件（95%予測区間上限：4 件）へと微かな増加に留まり，降水量の変化に対する影響差が定量的に示された．

以上より，本モデルは突発的な発生数の増加に対して過小評価を示す場合があるものの，都道府県ごとの地域特性や災害種別ごとの傾向を一定程度反映できていると考えられる．

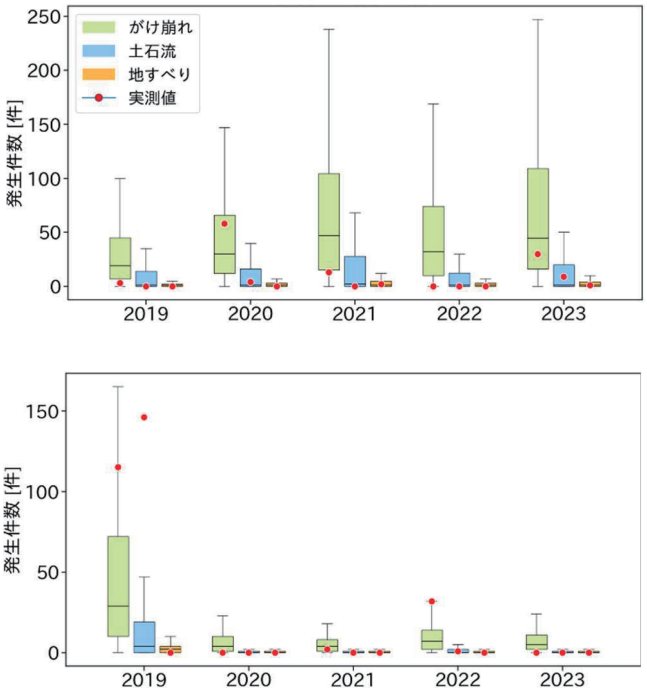


図-4 検証期間中における災害種別発生数の実測値と予測分布（上：福岡県，下：宮城県）

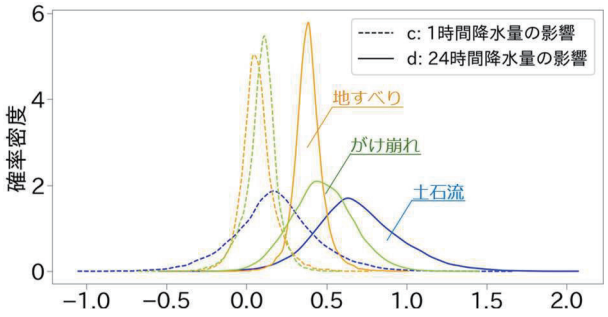


図-5 宮城県の降水量パラメータの事後分布

5. がけ崩れ，土石流および地すべりの推定パラメータに関する詳細分析

推定されたパラメータe^{a_j}について，土石流とがけ崩れではほとんどの都道府県で増加傾向を示した一方で，島根県と佐賀県では減少傾向を示した．これらの都道府県では過去に大規模な災害が発生していることが挙げられる（佐賀県 1990 年[7]，島根県 1983 年[8]）．本モデルの対象期間において，こうした記録的な突出値が学習期間の前半に含まれることで，統計上のトレンドが相対的に右肩下がりとして検出さ

れたと考えられる．

地すべりでは，他の二災害とは異なる傾向を取った．まず，発生しやすさを示すパラメータe^{b_j}は，新潟県，富山県，山形県などで高い値を示した．これらの地域では地すべりが発生しやすい地質が分布していると指摘されており[9]，既往の研究と高い整合性を示した．また，新潟県，富山県，山形県では，パラメータe^{a_j}も正の値を示した．これは，もともと不安定な地質構造を持つ地域において，近年の降雨パターンの変化や積雪・融雪状況の変化が，地すべりの活動性をさらに高めている可能性を示唆している．一方で，石川県や鳥取県のようにe^{b_j}が大きい値を取りながら，e^{a_j}が明確な負の値を示す地域が確認された．これは実データとも一致する結果である．この要因としては，これら両県において観測体制の変化や地すべり対策事業の効果などが可能性として挙げられる．また，もともと地すべり地形が少なく発生が稀な都道府県においては，e^{a_j}が負の値を示す傾向があった．これは，発生確率が極めて低い状態が継続していることをモデルが減少傾向として捉えている，あるいは微少な発生報告の変動が統計的に負のトレンドとして検出されたものと考えられる．

総じて，一部の都道府県において信用区間が広範となった主な要因は，対象期間内の土砂災害発生数が少なく，推定に利用できる情報量が限られていたためである．ベイズ推論では，発生件数が少ない地域ほどパラメータ推定の不確実性が大きくなる傾向がある．一方，ハッチングを付した地域では，年次変化を表す係数が統計的に有意であり，土砂災害発生数の経年的な増加または減少の傾向が示された．本モデルは，今後のデータ蓄積に伴う推定精度の向上が期待され，広域的な土砂災害リスク評価の基盤となることが期待される．

6. 結論

本研究では，1983 年から 2023 年における都道府県別の土砂災害発生傾向を分析するため，降水量および年次を説明変数とする階層ベイズモデルを構築した． モデルの妥当性について，検証期間における全国の実測値と予測値ですべての災害種別において$p > 0.05$であったことから，予測モデルは実測値の分布を適切に再現していた．

パラメータ推定の結果，増加傾向を示す都道府県と減少傾向を示す都道府県が明確に分類され，今後のリスク変化の地域差が明らかとなった．年次経過による発生数の変化を表すパラメータe^{a_j}について，がけ崩れにおいて，統計的に有意に増加傾向であったのは 21 県，土石流においては 11 県が有意に増加傾向であった．地すべりでは 11 県で統計的に有意に減少するという結果が得られた．

本研究で得られた成果から，階層ベイズモデルを用いることで，都道府県ごとの土砂災害発生特性や降水量に対する感度といった地域差を定量的に評価できることを示した．これらの知見は，地域特性を考慮した土砂災害対策や防災計画の立案に資することが期待される．また，本モデルは，将来の降水量を入力することで土砂災害発生数を確率的に推定できるため，被害予測や災害リスク評価に加え，地域特性を反映した土砂災害警戒基準線（CL 線）の設定や，災害発生事例が少ない地域における CL 線設定への応用も期待される．これにより，より適切な防災情報の提供や住民の避難判断支援への活用が期待される．

参考文献

1) 桑折奎吾，劉ウェン，丸山喜久：機械学習に基づく豪雨による土砂崩壊発生地点の予測と説明可能 AI による予測結果の分析，AI・データサイエンス論文集 3 巻 J2 号，2022.

2) 猪俣雄太・中田知沙・山口浩和：階層ベイズモデルによる林業労働災害の発生に及ぼす要因，森林利用学会誌 38(1)，pp. 13～22，2023.

3) 廣井悠：階層ベイズモデルを用いた地震火災の出火件数予測手法とその応用，地域安全学会論文集，27 巻，pp. 303-311，2015.

4) 砂防・地すべり技術センター，土砂災害の実態，1983-2023.

5) 総務省統計局，都道県地形・傾斜度別面積，https://www.stat.go.jp/data/nenkan/65nenkan/01.html.

6) 気象庁：過去の気象データ，1983～2023.

7) 藤原治，柳田誠，清水長正，三箇智二，佐々木俊法：日本列島における地すべり地形の分布・特徴，日本地すべり学会誌 41 巻 4 号，pp. 335-344，2004.

8) 砂防・地すべり技術センター，土砂災害の実態，1990.

9) 砂防・地すべり技術センター，土砂災害の実態，1983.

（2026.5.29　受付）