

# GIS を用いた地名と 自然災害の関係性評価

1200020 石川恵大<sup>1</sup>

<sup>1</sup>高知工科大学 システム工学群 建築・都市デザイン専攻

E-mail 200020i@ugs.kochi-tech.ac.jp

日本は古くから災害が多発していたことから、災害を示唆するような災害地名が多くあるなか、時代の移り変わりの中でまったく別の地名に変更されることがある。本研究は改めて地名と自然災害との関連性を評価することで地名の防災への利用を促進することを目標とし、対象地域の地名と災害箇所の関係性を考察する。分析の際には地名はその「読み」に着目し、災害箇所は機械学習にて抽出を行った。

四国全域を対象に、CS 立体図を用いた機械学習により土砂災害を抽出することができたが、災害地名と機械学習による土砂災害の抽出箇所との間に明確な関係性を確認することはできなかつた。今後は特徴量や分類方法の見直し等により土砂災害抽出の精度を向上させたいと、地名との関係性を評価することが必要であると考えられる。

**Key Words:** 災害地名, CS 立体図, 機械学習

## 1. 序論

### 1-1. 背景

地名には大きく分けて「自然地名」と「文化地名」があり、その土地の地形、植生、方位などの様々な要素が由来となっている。その中でも日本は古くから災害が多発していたことから、災害を示唆するような地名があり、一般的に「災害地名」と呼ばれている。例として、高知県中土佐町久礼や徳島県三好市大歩危などがあげられる。前者の久礼は古語の「抉る」が転じたもので、川の屈曲を表し、洪水による氾濫地帯や土砂災害の起こる傾斜地が確認されることが多い。後者の歩危は古語の「ホキ」が転じたもので川崖、山崖を表し、急斜面の断崖地や地すべり崩壊地、川岸の崖で水流が当たる場所にみられる。実例として、それぞれ 1981 年 6 月に斜面崩壊、1987 年 7 月に地

すべりが確認されている。東北大の花岡<sup>1)</sup>は、先の東日本大震災の被災地でもある岩手県、宮城県を対象に、地名語尾と災害リスク、土地条件を比較する研究を行い、地名語尾と自然災害リスクとの間に一定の関係があることを明らかにした。しかし、地名の中には長期間にわたり、地域で共有されるものもある中で、時代の移り変わりの中でまったく別の地名に変更されることがあり、本来の地名の意味を解釈することが困難になってきている。<sup>1)</sup>

### 1-2. 目的

本研究は背景で挙げた、①災害地名の存在②地名と自然特性の関連性③地域イメージを重視した地名の蔓延の 3 点を踏まえ、改めて地名と自然災害との関連性を評価することでその信ぴょう性を再認識し、地名の防災への利用を促進することを目標とし、対

象地域の地名と災害箇所の関係性を考察する。災害の指標に地名を組み込むことで住民になじみやすい対策が可能となり、地域住民の防災への主体性向上につながると考える。

### 1-3. 本研究の概要と独自性

本研究は、四国全域を対象とし、機械学習により抽出された災害地の分布と四国の地名との関係性を分析する。分析の際には地名は字、通称の「読み」に着目する。小川<sup>2)</sup>によると、地名は元来、話し言葉により伝承してきたものであり、その本来の意味を解釈するためには地名の読みによる解釈が必要であるとされている。また、地名の小字・俗称は小地域の土地の特徴が最もよく表現されており、自然災害との関係性を評価するにあたって今回の分析の対象とする。分析に用いる災害箇所は機械学習にて抽出する。機械学習を用いることで現在確認されている分布に加え、未確認の箇所や災害が発生するポテンシャルのある地域も抽出することができる。従来の等高線などを用いた地形判読による災害箇所の抽出は専門性が高く、困難であったが、長野県林業総合センターが考案した CS 立体図により地形判読を容易に行うことができるようになった。本研究はこの CS 立体図のもととなる、標高、傾斜、曲率の組み合わせが災害箇所抽出に特徴量として適していると考え、機械学習を行った。

## 2. 使用データ

### 2-1. 地名データ

地名データは国土地理院が発行する数値地図（国土数値情報）を使用した。収録データはポイントデータで、都市域では縮尺 2500 分 1 相当以上、その他地域では縮尺 25000 分 1 相当以上の位置精度となっている。また、収録されている地名は、大字・町目、字・通称レベルである。1-3 で述べたように、小字・俗称が最もその土地の特徴を表すということから、本データから字・通称のみを取り出した。

### 2-2. 機械学習に用いたデータ

機械学習に用いたデータは以下のとおりである。

#### 1) 教師データ

・地すべり：1:50,000 地すべり地形分布図（取得元：防災科学技術研究所）

・土石流：国土数値情報 土砂災害危険箇所（取得元：国土交通省 精度：1/25000 程度）

・急傾斜地崩壊：国土数値情報 土砂災害危険箇所（取得元：国土交通省 精度：1/25000 程度）

#### 2) 特徴量

<CS 立体図>

・数値標高モデル（取得元：国土地理院）

・傾斜データ (QGIS 内の SAGA により DEM から算出)

・曲率データ (QGIS 内の SAGA により DEM から算出)

<追加データ>

・傾斜方位データ (QGIS 内の地形解析により DEM から算出)

・20 万分の 1 日本シームレス地質図（取得元：産業技術総合研究所 地質調査総合センター）

## 3. 機械学習による災害箇所の抽出

### 3-1. 抽出手法

四国スケール対象に機械学習により地すべり地形、土石流、急傾斜崩壊地の抽出を行った。今回教師データとして用いた地すべり地形分布図は比較的大規模な地すべり地の分布を空中写真から判読したものであるのに対し、土石流、急傾斜崩壊地は実際に発災した分布ではなく、その危険性のある地域において定められた条件に適合するか調査した結果から指定されたものとなっている。そのため、機械学習後の結果としては地すべりの場合、未確認のものを含めた、実際に地すべりが発生したと考えられる場所の分布が抽出され、土石流、急傾斜崩壊に関しては発災の危険性がある場所の分布が抽出されることが考えられる。機械学習 Python の機械学習ライブラリ scikit-learn を使い、サポートベクターマシンで災害箇所をクラス 1、起きていない箇所をクラス 0 とする分類を行った。分類の際の特徴量には、標高デー

た、傾斜データ、曲率データのほかに傾斜方位と地質データを用いた。解析メッシュサイズは災害箇所のパolygonサイズと、地名との関係性評価を考慮して、地すべり地形は500mメッシュに、土石流、急斜面崩壊地は250mメッシュに設定した。この際、標高、傾斜、曲率、傾斜方位に関しては一つのメッシュに平均値を、地質に関しては最頻値を格納させた。災害箇所抽出フローを図-1に示す。

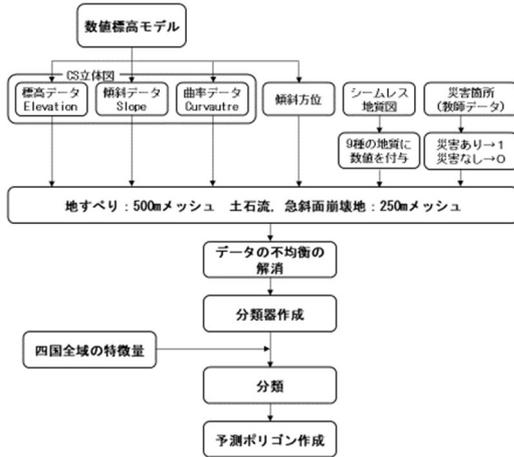


図-1 災害箇所抽出フロー

また、分類器のトレーニングデータを取得した範囲を図-2に示す。



図-2 トレーニングデータ取得範囲

### 3-2. 分類結果と考察

4通りの特徴量の組み合わせによって学習した分類器を用いて四国全域の特徴量について分類を行った。その後、適合率(precision)及び再現率(recall)を用いて分類精度の評価を行った(表-1, 表-2)。ここで、適合率(precision)とは分類結果がクラス1と予測されたサンプルのうち正解したサンプルの割合である。また、再現率(recall)は実際にクラス

1と分類されたサンプルの内正解したサンプルの割合である。適合率、再現率は以下の式であらわされる。

$$\text{precision} = \frac{TP}{TP + FP}$$

$$\text{recall} = \frac{TP}{TP + FN}$$

表-1 TP(真陽性), TN(真陰性), FP(偽陽性), FN(偽陰性)は以下の混合行列により定義される。

混合行列		予測	
		Negative	Positive
正解	Negative	TN	FP
	Positive	FN	TP

表-2 分類精度の評価結果

	地すべり		土石流		急傾斜崩壊地	
	precision(%)	recall(%)	precision(%)	recall(%)	precision(%)	recall(%)
ver01	98.3	71.9	99.8	67.2	99.4	53.0
ver02	98.4	71.3	99.8	68.5	99.4	53.7
ver03	98.4	71.7	99.8	67.6	99.4	53.6
ver04	98.3	71.4	99.8	68.8	99.4	53.7

ver01 特徴量: 標高, 傾斜, 曲率

ver02 特徴量: 標高, 傾斜, 曲率, 傾斜方位

ver03 特徴量: 標高, 傾斜, 曲率, 地質

ver04 特徴量: 標高, 傾斜, 曲率, 傾斜方位, 地質

本研究では、未確認の分布や発災のポテンシャルのある個所の抽出を考慮し、4通りの特徴量の組み合わせの内、適合率の最も高いもののうち、再現率が高いものを精度の高いものとし、採用した。地すべり地形はver03, 土石流はver04を、急傾斜崩壊地はver04を採用した。精度評価として、適合率、再現率の良いものを採用して学習させたものの、CS立体図上では平野部など、明らかに起きていないと判読される箇所でクラス1と分類されている箇所が多く見受けられた。一方、未確認分布も抽出されていることが教師データでは災害箇所とされていないが、CS立体図からは災害が発生した可能性があるかと判断できる場所も確認された。今後は土

地利用や植生分布等を考慮した分類により、精度向上を図ることが必要だと考える。分類結果から災害箇所と判断されたメッシュを以下に示す。

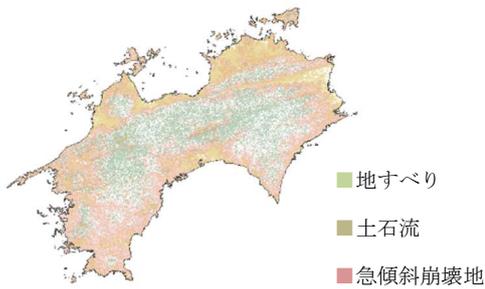


図 - 3 機械学習による災害箇所の分類結果

## 4. 自然災害と地名との関係性評価

### 4-1. 評価手法

地名と災害箇所との関係性を2つの評価方法を用いて評価した。小川<sup>2)</sup>の著書から選定し、評価に用いた災害地名読みは四国における土砂災害を示唆する地名読みを表-2に示す。

評価手法1) 災害地名読みが含まれる地名と含まれない地名ポイントからそれぞれ1kmのバッファを生成し、そのバッファに含まれる機械学習により抽出された災害箇所のメッシュ数の平均を計算する。

評価手法2) 災害地名読みが含まれる地名ポイントが機械学習によって抽出された災害箇所メッシュに含まれている割合。

表 - 3 四国域の土砂災害地名読み

ほき	ほけ	ほげ	ほか	ほが	ほか	ふき	あず
ぼーが	され	さる	くい	くえ	つえ	くれ	あづ
かけ	かげ	がけ	かぎ	かき	しし	けた	うめ

## 5. 結果と考察

表-3, 表-4に2つの評価法に基づいた分析結果を示す。

表 - 4 評価手法1の結果

	地すべり(個)	土石流(個)	急傾斜崩壊地(個)
災害地名読みあり	2.81	21.7	26.6
災害地名読みなし	2.04	25.3	25.7

表 - 5 評価手法2の結果

	四国全域の災害地名数	ポリゴン内の災害地名数	割合(%)
地すべり	443	151	34.1
土石流	443	387	87.4
急傾斜崩壊地	443	394	88.9

機械学習により抽出した土砂災害分布と四国域の災害地名との関連性を2つの評価手法を用いて考察する。評価1からは災害地名読みが含まれている地名もそうではない地名もその周辺の機械学習で抽出された土砂災害のメッシュ数にあまり大きな差は見られなかった。これは、教師データでは災害発生箇所とされていないが、機械学習では災害発生箇所と分類される場所が多かったためであると考えられる。一方、評価2からは四国全域の災害地名読みが含まれる地名が抽出された土砂災害分布内に一定数含まれることがわかった。

## 6. まとめ

本研究では地名と自然災害の関係性を評価し、改めて地名の防災への有用性を認識することを目的としたが、機械学習によって抽出された災害分布と地名との間に明確な関連性は確認されなかった。しかし、古くから伝わる地名はその土地の文化や自然を表しているものが多いということは事実であり、今後の利用促進のためさらなる研究と情報発信が必要である。

## 7. 参考文献

- 1) 花岡和聖：小地域地名の語尾と自然災害リスクの関係性評価
- 2) 小川豊：危険地帯がわかる地名
- 3) 長野県林業総合センター：CS立体図の開発
- 4) CS立体図を用いた地形判読：

<https://www.slideshare.net/osgeojapan/csfoos4g-2017-tokyo>