

# 機械学習を用いた斜面危険度評価手法の高度化

環境防災研究室

五十嵐 大輝

指導教員

大塚 悟, 福元 豊

## 1. はじめに

斜面の広域危険度分析は、災害因子に対して対象地点のポイントワイズな統計的評価や AI 評価が行われているが、土塊が一体化して挙動する斜面崩壊の実現象と大きく乖離する問題点を有する。また、近年の機械学習を用いた評価解析は、多様なデータを活用するため、学習範囲と適用範囲に制限の生じる固有の課題があり、応用に関する知見が十分に整備されていない。本研究は斜面の広域危険度評価の高度化を目的に、本研究室で開発してきた機械学習による解析手法を用いて、実務での適用範囲の拡大に向けた AI 評価モデルの簡素化と高精度化を目指した検討を行った。

## 2. 解析範囲と地形情報データの構築

本研究では、平成 23 年 7 月新潟・福島豪雨において斜面災害調査が実施された魚野川流域の崩壊地データを使用した。流域全体では膨大なデータ数となるため、流域内から Area\_01 (5,000×5,000m の矩形部) を解析領域として選定した。また他地域への適用性を検討するため、標高や総雨量、地質状況の異なる Area\_02~04 (5,000×5,000m の矩形部) を合わせて選定した。

崩壊予測に関する素因データとして、災害前の 1mDEM から、ArcGIS を用いて、(1) 標高を算出し、標高データより(2) 傾斜角、(3) 曲率、(4) 断面曲率、(5) 平面曲率、(6) 起伏量、(7) 累積流量、(8) 地上開度、(9) 地下開度、(10) 尾根谷度の地形量を作成した。また、崩壊規模の大小に対して、効果的な予測実現のため、スケールの異なる地形情報として、1m メッシュを中心として 5×5 メッシュ分の移動平均値 (移動平均 5m)、10×10 メッシュ分の移動平均値 (移動平均 10m) の地形量も作成した。作成した地形量間の相関係数を算出した結果、複数の地形量に対して、高い相関が確認された。

- ① 傾斜角と起伏量に高い正の相関あり (r=0.95)
- ② 曲率と断面曲率に高い正の相関あり (r=0.86)

- ③ 曲率と平面曲率に高い負の相関あり (r=-0.87)
- ④ 地上開度と尾根谷度に高い正の相関あり (r=0.81)
- ⑤ 地下開度と尾根谷度に高い負の相関あり (r=-0.78)

## 3. 機械学習による斜面崩壊危険度評価

### 3.1 モデル評価手法

本解析では、Area-01 におけるデータを基に機械学習のバギング木による崩壊予測を実施した。応答変数を災害の崩壊・非崩壊の事象とし、説明変数は地形量 (2)~(10)とした。モデルの分類性能の評価には図 1 の混同行列から算出される真陽性率、真陰性率、g-mean を用いた。また、これらの指標ではモデルの判定結果から得られる危険度評価値 (0~1 の値をとる) そのものを直接評価できないことから、0.4~0.6 と曖昧な判定を受けたデータ数を用いた「曖昧度」を定義し、予測のメリハリ度合いを評価した。

$$\text{真陽性率} = \frac{TP}{TP + FN}$$

$$\text{真陰性率} = \frac{TN}{TN + FP}$$

$$g - \text{mean} = \sqrt{\text{真陽性率} \times \text{真陰性率}}$$

$$\text{曖昧度} = \frac{\text{スコア確率が0.4~0.6のデータ数}}{\text{データ総数}}$$

		予測結果	
		崩壊	非崩壊
正解	崩壊	TP	FN
	非崩壊	FP	TN

図1 混同行列

### 3.2 重要度と相関係数による特徴量選択

本研究では、従来の解析で使用していた 9 つの地形量に対し、いくつかの変数の削減を行った。削減する地形量の選定では従来解析より得られた重要度と相関係数を用いた。重要度はモデルに入力した地形量が予測評価にどの程度影響を与えているのかを表す。図 2 に示す重要度の内、高い値を示した上位の項目は傾斜角、曲率、地上開度、地下開度、尾根谷度であった。また、これらの相関性を見ると、地上開度と尾根谷度で

0.86, 地下開度と尾根谷度で-0.78 と高い相関性を確認した. そこで, 従来の解析モデル (Case1) と重要度の上位 5 項目だけを用いた新モデル (Case2), 重要度が高く, 相関係数が高い変数同士のうち片方または両方を削除したモデル (Case3~7) を作成し, 性能を比較した. 結果を表 1 に示す. 従来モデルの真陽性率は 89.5% に対し, 入力地形量を重要度上位 5 項目に削減したモデルでの真陽性率は 85.2% となり, 4 ポイントの低下に留まった. また, 相関係数の高い項目の中で, 尾根谷度を除いた Case3 は全性能指標において低下がみられなかった. この結果から, 全地形量の入力不要で, 重要度の高い項目の入力と相関の高い項目の削減 (多重共線性の排除) で高精度の危険度予測が可能であり, 計算負荷の軽減に寄与することを確認した.

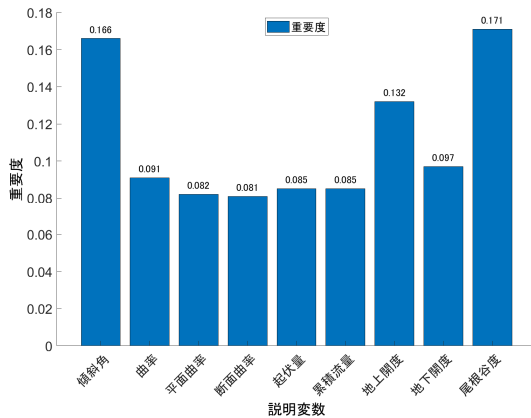


図2 各地形量における重要度

表1 特徴量選択での性能指標の変化

モデル	真陽性率	真陰性率	g-mean	曖昧度	
				崩壊箇所	非崩壊箇所
Case1	89.5%	77.4%	83.3%	17.1%	21.1%
Case2	85.2%	74.1%	79.5%	20.9%	22.3%
Case3	85.4%	74.3%	79.7%	20.1%	22.2%
Case4	77.9%	65.4%	71.4%	28.6%	29.4%
Case5	76.6%	65.6%	70.9%	28.8%	29.5%
Case6	76.3%	67.6%	71.8%	25.8%	26.2%
Case7	65.8%	52.8%	58.9%	45.7%	43.0%

### 3.3 スケール情報の選択による性能変化

本節ではスケール情報の選択が危険度予測に与える影響を調査した. 1m メッシュ, 移動平均 5m, 移動平均 10m の組み合わせから全 7 ケースについて解析した (表 2). 入力地形量は前解析で得られた結果に基づき, 傾斜角, 曲率, 地上開度, 地下開度の 4 項目とした. 解析結果を図 3 に示す. 真陽性率を比較すると, Case2 で最大 86.5%, Case7 で最低の 68.6% となった. 高解像度の 1m メッシュによる地形情報は, 詳細な微地形表現や実現象に即した崩壊形状の把握を可能にする一方, ノイズによる過学習の影響を受

けることがわかった. 移動平均 5m と 10m のみの入力が, 1m メッシュも入力したケースよりも予測性能が高い結果を得た.

表2 作成した学習データの概要

ケース	入力するスケール	項目数
Case1 (12 項目)	1m メッシュ	1m メッシュ: 4 項目
	移動平均 5m	移動平均 5m: 4 項目
	移動平均 10m	移動平均 10m: 4 項目
Case2 (8 項目)	移動平均 5m	移動平均 5m: 4 項目
	移動平均 10m	移動平均 10m: 4 項目
Case3 (8 項目)	1m メッシュ	1m メッシュ: 4 項目
	移動平均 10m	移動平均 10m: 4 項目
Case4 (8 項目)	1m メッシュ	1m メッシュ: 4 項目
	移動平均 5m	移動平均 5m: 4 項目
Case5 (4 項目)	移動平均 10m	移動平均 10m: 4 項目
Case6 (4 項目)	移動平均 5m	移動平均 5m: 4 項目
Case7 (4 項目)	1m メッシュ	1m メッシュ: 4 項目

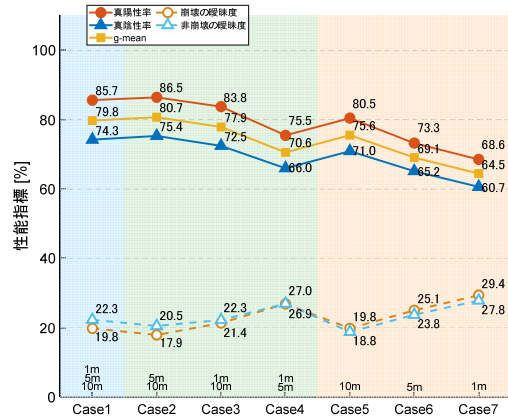


図3 スケール情報の組み合わせによる予測性能の比較

### 4. 他地域への適用性拡大の検討

本節では, Area-01 で高い予測精度が得られた条件を他地域に適用した場合について検討した. Area-02~04 の解析結果を表 3 に示す. Area-01 でのモデル条件では, 真陽性率は平均 40% となり, 十分な性能が得られなかった. ただし, エリアによって性能が異なり, Area-02,03 での真陽性率は平均 35% であるのに対し, Area-04 の真陽性率は平均 47% で Area-04 が 12 ポイント高いことを示した. これは, エリア間の地質の類似性に依存しているためと考える. Area-02 と Area-03 の地質が新第三紀で同じため, 真陽性率が同程度となり, Area-04 は Area-01 と同じ深成岩類であるため真陽性率が相対的に高くなったと考える.

他地域への適用において, Area-01 での検討と同様に, Area-02~04 での検討でも重要度上位 5 項目と相関係数による入力地形量の選定による性能変化の比較を行った. Area-02 の結果を図 4 に示す. Case3 では真陽性率は低いですが崩壊の曖昧度は低い結果を得た. また, Case5 では真陽性率は高いが崩壊の曖昧度は高い結果を示した. こ

れらは Area-03,04 でも共通してみられた予測性能である。曖昧度が低い Case3 は境界付近の予測をより明確に区分する「メリハリモデル」として有用である。また、Case5 は真陽性率が高いため、防災上の重要課題である崩壊見逃しを極力回避する観点から「安全評価モデル」として有用であると考えられる (図5)。

表3 他地域での検証結果

学習→テスト	モデル	真陽性率	真陰性率	g-mean	曖昧度	
					崩壊箇所	非崩壊箇所
01 → 02	1	35.9%	80.2%	53.7%	24.8%	17.0%
	2	33.7%	83.0%	52.9%	25.2%	16.3%
01 → 03	1	35.9%	77.5%	52.7%	25.8%	19.0%
	2	34.4%	78.4%	52.0%	23.8%	17.4%
01 → 04	1	48.2%	73.8%	59.7%	27.9%	21.9%
	2	45.3%	74.9%	58.2%	26.5%	19.4%

モデル1: 地形量選定モデル    モデル2: スケール情報選定モデル

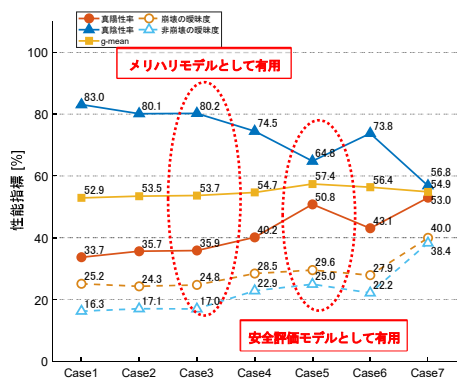


図4 Area-02 での特徴量選定による性能変化

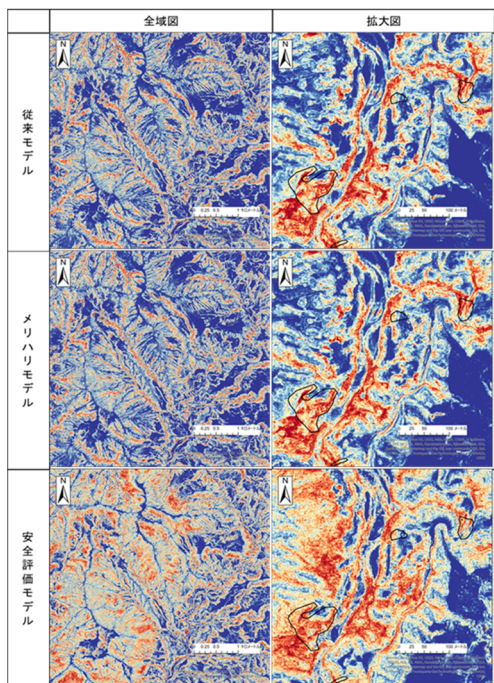


図5 提案する手法による危険度評価 (Area-02 検証)

## 5. まとめ

本研究では、斜面の広域危険度評価において、機械学習による解析モデルの高度化を目指した検討を行った。多様な地形量・スケール情報の中で、崩壊予測に寄与する情報を見出したことで、少ない入力情報により経済的かつ従来と同様に高精度な予測を実現できた。Area-01 で得られた評価モデルを他地域へ適用した結果、地質が類似する地域では予測性能が向上することを示した。今後は地形特性や降雨の地下浸透の大小などを反映した崩壊発生の素因データの開発、入力項目の選定等を行い、汎用性のある評価手法の検討を継続する必要がある。