

AI を用いた豪雨時の斜面災害の広域予測

長岡技術科学大学 非会員 ○片平耀款
長岡技術科学大学 正会員 大塚悟
長岡技術科学大学 正会員 福元豊

1. はじめに

近年、日本各地において集中豪雨による各種土砂災害が多数発生している。最近 10 年余では、平成 23 年 7 月新潟・福島豪雨、平成 26 年 8 月広島豪雨、平成 30 年 7 月豪雨において発生した斜面災害により、多くの尊い命が失われた。豪雨発生時の環境や崩壊の状態によって被害の形態は様々であるが、豪雨による斜面崩壊が多くの人命に関わるという点は共通の事実である。集中豪雨が多発する我が国において、斜面崩壊の箇所をあらかじめ予測することは、防災の観点から非常に重要であると考えられる。

また、最近になり AI を用いた崩壊に関する研究が行われてきたが、これまで研究が行われてきたものは、メッシュにおける属性（標高、傾斜、地質）などを用いて、同メッシュの崩壊・非崩壊との相関を解析するものである。そのため、メッシュ単位の解析を実施しており、隣接するメッシュとの影響をはじめ、斜面安定解析で考慮するすべり土塊全体の影響を考慮する解析は GIS を用いた広域危険度評価では実施されていない。本研究では、広域危険度で評価に用いた 1m メッシュのデータに、周囲 5m の情報及び 10m の巨視的な地形情報を加えることで、広域的なデータを入力したと考え、そのデータについて予測を行い、危険度評価における有効性について検討する。

2. 調査地域と資料概要

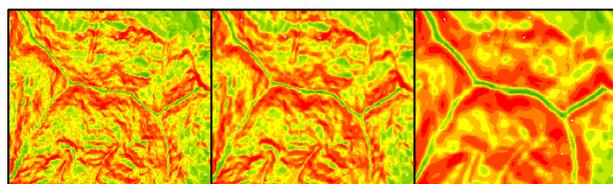
今回用いたデータは、平成 23 年 7 月に発生した新潟・福島豪雨で、崩壊の有無の調査が行われた図-1 の黄色範囲のものである。対象地域には、新潟県魚沼市、南魚沼市の主な地域と長岡市、小千谷市の一部が含まれている。また、1m メッシュで以下の範囲内の領域をすべて扱うには、データ数が非常に膨大であるため、図-1 の範囲からさらに対象地域を絞った赤



図-1. 調査地域及び対象範囲

色の範囲を今回の対象範囲とした。調査地域での崩壊面積は、最小が 3.5m²、最大が 24,319m² となっており、1m メッシュデータの適用が可能である。赤の対象範囲では、範囲が狭いため、降雨量に変化がみられず、予測子として AI の予測に使用しても意味がなくなる。そのため、今回は降雨量を除いた定性的なデータでの予測を行うことにした。

本研究では、応答変数を災害時の崩壊・非崩壊の事象とし、予測子変数を 1.標高、2.傾斜角、3.曲率、4.断面曲率、5.平面曲率、6.起伏量、7.累積流量の 7 項目とした。また、広域範囲を含んだデータでは、1~7 に加えて、1m メッシュのデータと同項目の 5m メッシュのデータを 7 項目、10m メッシュのデータを 7 項目含んだ計 21 項目を入力変数とした。上記のデータについては、ArcGIS(ESRI 社)によって、標高から導き出すことができる。また、今回用いた DEM データは、国土交通省北陸地方整備局湯沢砂防事務所が 2010 年 7 月に作成したものを用いている。



(A) (B) (C)

図-2. 傾斜角のデータ ((A) 1m, (B) 5m, (C) 10m)

3. 方法

ArcGIS で数値化したデータを、MATLAB の機械学習ツールを用いて予測を行った。今回予測に用いたのは機械学習の分類法である決定木、KNN、バギング木の三種類の方法である。決定木は、データがある条件に当てはまるか当てはまらないかで分類する方法。KNN は学習したデータに近いものを予測の結果とする方法。バギング木は、多数の弱学習器から多数決をとったものを出力する方法である。データの入力方法は、数値を直接代入する方式をとった。調査地域では、河川や池・湖などの調査不能な範囲を除いて、

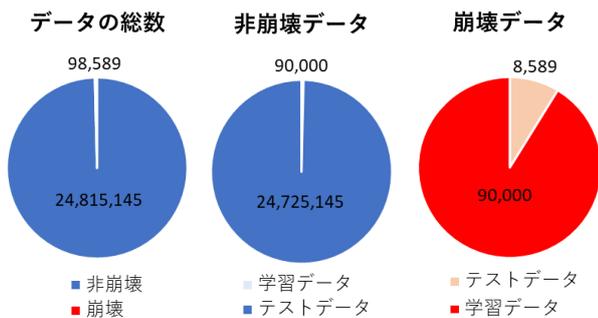


図-3. 各データ数の内訳

2,4913,734 個のデータを得ることができた。そのうち崩壊のデータは、98,589 個であった。崩壊データ数のうち 90,000 個を学習データとし、残りの 8,589 個をテストデータとした。過学習を起こさないように、非崩壊の学習データをランダムに 90,000 個抽出し、崩壊の学習データの個数と同数にした。計 180,000 個の学習データのうち 25% を検証データとして、ホールドアウト検証を行い、データが過学習を起こしていないことを確認した。

4. 解析結果

解析結果を表-1 に示す。崩壊予測としては、決定木において正答率・真陰性率が 1m メッシュより低くなってしまったが、バギング木・KNN においてはすべての値で 1m メッシュのみのデータより、広域的な情報を入れたもので精度が向上することが確認できた。決定木においても、実際の崩壊と予測の崩壊の正答率を示す真陽性率において、1m メッシュのみのデータを大きく上回っている。これは、1m メッシュのみの微地形の特徴だけでは予測することが困難だったデータを、5m および 10m の特性長の傾斜角や曲率を取り入れて巨視的情報を含んだ解析を行うことで予測することができたことを示している。これにより、精度が向上したのだと考えられる。また、非崩壊の学習データはランダム抽出を行ったが、約 2,400 万個のデータから 9 万個を選んでいる。そのため、1m のみのデータだと全体をとらえきれず予測のデータに適したデータのみを抽出する可能性があるが、5m, 10m の情報を入れることで、その可能性を低くすることもできると考えられる。ただ、今回の結果では偽陰性率・偽陽性率ともに高くなっているため、この結果自体を今後の予測に用いることは難しい。分類の精度が低くなってしまったのは、データの代入方法に課題があると考えられる。

表-1. 手法ごとの解析結果

(%)	決定木		バギング木		KNN	
	1m	1m,5m,10m	1m	1m,5m,10m	1m	1m,5m,10m
正解率	70.6	64.5	63.1	77.9	54.0	61.8
精度	62.3	64.6	60.7	78.3	54.1	62.9
真陽性率	48.6	64.8	57.1	79.7	53.4	57.8
真陰性率	70.6	64.5	63.1	77.9	54.6	65.8
偽陰性率	51.4	35.2	42.9	20.3	46.6	42.2
偽陽性率	29.4	35.5	36.9	22.1	45.4	34.2

表-2. 混同行列

		予測のクラス	
		崩壊(1)	非崩壊(0)
実際のクラス	崩壊(1)	TP	FN
	非崩壊(0)	FP	TN

$$\text{正解率} = (TP+TN) / (TP+TN+FP+FN) * 100$$

$$\text{精度} = TP / (TP+FP) * 100, \text{ 真陽性率} = TP / (TP+FN) * 100$$

$$\text{真陰性率} = TN / (TN+FP) * 100, \text{ 偽陰性率} = FN / (TP+FN) * 100$$

$$\text{偽陽性率} = FP / (TN+FP) * 100$$

今回の方法では、GIS にて数値化されたデータをそのまま機械学習に用いた。機械学習では前処理でデータのクラス分けを行うことで精度が向上することがある。今回は、クラス分けにより正確な値が失われる可能性を危惧し、行わなかったが今後検討していきたい。

5. 今後の展望

今後の展望として、まず予測したデータを GIS 上に戻し、既存の崩壊データと予測データとの地図上での比較を行いたいと考えている。それにより、どのような箇所で予測が外れたか、また今後どのようにすれば改善することができるか考えることができる。

また、1mDEM データの入手は限定された範囲のデータが多い。今回用いたデータも豪雨で被害を受けた範囲のものである。今後は、簡単に入手可能でより広範囲で予測を行うために、5m のデータに 10m のデータを入れて予測を行いたいと考えている。

6. 参考文献

呉 振宇ら(2020) : Landslide susceptibility mapping based on artificial neural network (ANN) and Geographic Information System (GIS), 地盤工学会北海道支部技術報告集, 295~304.

7. 謝辞

データを提供頂いた国土交通省北陸地方整備局湯川砂防事務所および、技術的な助言を頂いた北海道大学磯部公一先生に謝意を表します。また、本研究は東京電力ホールディングス株式会社の支援により実施しました。