

地震動継続時間を考慮した機械学習に基づく液状化の広域予測手法の構築

Development of a Wide-Area Prediction Method for Liquefaction Based on Machine Learning
Considering the Duration of Earthquake Motion

20WM1321 中山 洋斗
Hirotō Nakayama

指導教員 丸山 喜久
劉 ウェン

SYNOPSIS

In the 2011 Tohoku Japan Earthquake, liquefaction occurred in various areas in the eastern Japan. In the current liquefaction prediction, geotechnical survey data is employed, and it gives accurate results. However, it is difficult to apply this method to a wide area. Therefore, it is expected to establish a method which can easily evaluate the occurrence of liquefaction extensively. In this study, we aim to construct a liquefaction prediction model using support vector machine and random forests. Based on the results of previous studies, the duration of seismic motion, which affects the occurrence of liquefaction, is newly added as an explanatory variable. In addition, the models are grouped according to the tendency of liquefaction occurrence based on the Japan Engineering Geomorphologic Classification Map.

1. 研究背景と目的

2011年の東北地方太平洋沖地震では、青森県から神奈川県までの広範囲に及び液状化が発生した¹⁾。液状化が発生すると、一般の構造物のほかにも橋梁やライフライン施設にも甚大な被害を与える。都市ガスのリアルタイム地震防災システムに液状化予測機能が搭載されている²⁾ことを踏まえても、液状化が発生した地点を迅速に予測することは発災直後の緊急対応に大きく貢献できるものと期待される。地点ごとの液状化の危険度評価は、地盤調査データから液状化指数(PL値)³⁾によって実施されることが多く、高精度な評価が可能である。しかし、詳細な地盤調査データを広域に取得することは困難であることから、道路や水道、都市ガスのようなライフラインネットワークの被害予測には、簡便に面的な評価を行える手法の確立が望まれる。例えば松岡ら⁴⁾は、微地形区分⁵⁾ごとに液状化発生危険度を評価する方法を構築している。また、桑原・松岡⁶⁾は、機械学習を用いて日本全国の液状化危険度を評価する手法を構築した。柳瀬・丸山⁷⁾は若松・先名⁸⁾が取りまとめた東北地方太平洋沖地震における液状化発生地点データを利用し、千葉県内の液状化地点と地形的特徴の因果関係を共分散構造解析で評価し、計測震度と微地形区分のほかに傾斜角も液状化の発生確率に影響があることを示した。さらに、この分析結果に基づき、サポートベクターマシン(SVM)による液状化発生の予測モデルを作成した。

既往研究⁷⁾では、SVMに入力する地震動に関する変数として計測震度だけを用いている。一方、液状化発生率と地震動の継続時間には相関関係があると評価されており、これを用いた液状化発生率の予測法も提案されている⁹⁾。

そこで本研究では、広域性及び即時性を指向した液状化発生予測モデルの構築を目指し、地震動の継続時間を考慮した機械学習を実行する。これによって、既往研究⁷⁾よりも精度が高い予測モデルの構築を目標とする。

2. 本研究で用いたデータ

本研究では、柳瀬・丸山⁷⁾と同様に、東北地方太平洋沖地震の際の千葉県を対象とする。東北地方太平洋沖地震で



図-1 千葉県の液状化地点⁸⁾

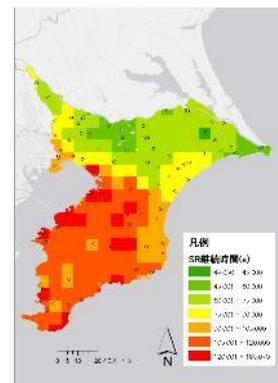


図-2 地震動継続時間の分布

は前震、余震を含め、複数の地震が発生しているが、液状化の発生には本震の影響が最も大きかったものと仮定する。本研究で使用したデータは、防災科学技術研究所の地震ハザードステーション(J-SHIS)の微地形区分、防災科学技術研究所の強震観測網(K-NET, KiK-net)の加速度記録、千葉県強震観測網の加速度記録、および産業総合技術研究所のQuiQuakeによる計測震度分布である。また、液状化発生地点には、図-1に示す若松・先名⁸⁾がまとめた5次メッシュごとに集計されたデータを使用する。加速度記録は地震動の継続時間を算出するために使用した。地震動の継続時間は、地動加速度の累積パワー曲線の占有時間から算出されるSR(Significant/Relative)継続時間⁹⁾とし、累積パワーが5~95%の範囲の時間を地震動の継続時間とした。これを観測点ごとに算出し、GIS上の内挿補間の手法の一つであるIDW法(Inverse Distance Weighted: 逆距離加重法)によって空間補間し、250mメッシュごとの地震動継続時間の推定値を算出した(図-2)。

3. 機械学習に基づく液状化発生予測モデル

本研究では教師あり学習を用いたパターン認識手法の一

表-1 精度評価のための混合行列

		実際のデータ	
		液状化あり	液状化なし
分析結果	液状化あり	TP	FP
		(True Positive)	(False Positive)
	液状化なし	FN	TN
		(False Negative)	(True Negative)

表-2 液状化の発生傾向に基づく微地形のグループ

区分	微地形区分	メッシュ数	液状化発生数	液状化率
グループ1	埋立地、干拓地、旧河道、自然堤防、河原	5860	1569	0.268
グループ2	扇状地、砂州・砂礫州	5671	212	0.037
グループ3	後背湿地、砂丘・砂丘間低地、三角州・海岸低地、砂丘	9564	1136	0.119
グループ4	谷底低地、砂礫質台地	12625	176	0.014
グループ5	山地、山麓地、丘陵、火山地、火山山麓地、火山性丘陵、岩石台地、ローム台地、岩礁	46092	661	0.014

つであるサポートベクターマシン (SVM) と、複数の決定木を用いて分類を行う機械学習アルゴリズムであるランダムフォレスト (RF) の2種類の機械学習を用いた。千葉県 の 79812 メッシュを学習データとテストデータに 7:3 に分割し、液状化発生予測モデルの構築と精度検証を行った。モデルの精度評価指標には、正解率 (Accuracy: *Acc*)、適合率 (Precision: *Pre*)、再現率 (Recall: *Rec*)、適合率と再現率の調和平均である F 値の4つの指標を用いる。各指標値は式(1)~(4)の通りであり、式中の TP, FP, FN, TN は表-1 の通りである。

$$Acc = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Pre = \frac{TP}{TP+FP} \quad (2)$$

$$Rec = \frac{TP}{TP+FN} \quad (3)$$

$$F = \frac{2 \cdot Pre \cdot Rec}{pre+Rec} \quad (4)$$

(1) サポートベクターマシン

サポートベクターマシン (SVM) は、教師あり学習を用いたパターン認識手法の一つである。基本的には2クラス分類のための手法とされており、認識性の優れた学習モデルである^{7),11)}。

本研究では、説明変数を地震動継続時間、計測震度、微地形区分としたモデル (SVM1) と説明変数を、計測震度、微地形区分としたモデル (SVM2)、既往研究^{5),6)}を基に液状化傾向から微地形を5つのグループに分けそれぞれモデルを作成した SVM3 の3パターンを検証した。SVM3 の各グループの微地形とそのメッシュ数を表-2 に示す通りである。液状化発生地点数と非液状化地点数が大きく異なるため、これらの比に基づき液状化発生地点に重みづけを行った。学習データで分類器の作成後、テストデータに対して SVM1 と SVM2, SVM3 に適用し、精度評価を行った。

SVM1 と SVM2 を比較すると、地震動の継続時間を加えることで液状化の見落としを大幅に減らすことができ、F 値が約 0.18 向上した。SVM1 と SVM3 では、液状化が多いグループ1 とグループ3 での精度がともに高くなった。その他のグループではやや指標に差があるものの、全体では

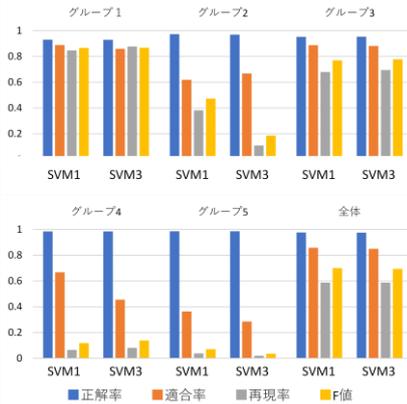


図-3 SVM1 と SVM3 の精度評価指標の比較

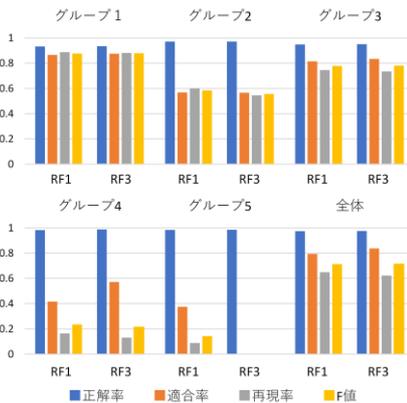


図-4 RF1 と RF3 の精度評価指標の比較

4 指標全てでほぼ同じ値となった (図-3)。

(2) ランダムフォレスト

ランダムフォレスト (RF) は複数の決定木を用いる集団学習法の一つである^{6),12)}。各決定木での予測結果を多数決することによって最終的な出力結果を得る。決定木学習とは、それぞれの接点 (ノード) でどちらに進むべきかを決定する分割関数を与え、最終的な出力結果であるクラスラベル *c* の事後確率 $p=c|v$ が定まるまで決定木を成長、分類させていく学習法である。

RF での SVM と同様に、説明変数を地震動継続時間、計測震度、微地形区分としたモデル (RF1) と、説明変数を計測震度、微地形区分としたモデル (RF2)、既往研究^{5),6)}を基に液状化傾向から微地形を5つのグループに分けそれぞれモデルを作成した RF3 の3パターンを検証した。RF1 と RF2 を比較すると、液状化の見落としを大幅が減り、正解率、適合率、再現率、F 値の全指標で精度が向上した。

RF1 と RF3 では精度指標値はグループごとに見ると、液状化の発生数が多いグループ1、グループ3 ではともに精度が非常に高くなった (図-4)。特にグループ1 は F 値が 0.879、またその他の指標も 0.86 以上となった。一方、グループ5 では、RF3 は液状化と予測した地点がなく、再現率は 0、適合率と F 値は計算できない。そのほかのグループについては、多少のばらつきはあるが、RF1 の方が多く液状化を予測する。全体では、RF1 と RF3 の F 値はほぼ同じとなり、再現率は RF1 のほうが高くなった。このことから RF1 の予測結果の方が良好であるといえる。

SVM1 と RF1 のテストデータに対する分類結果を図示し

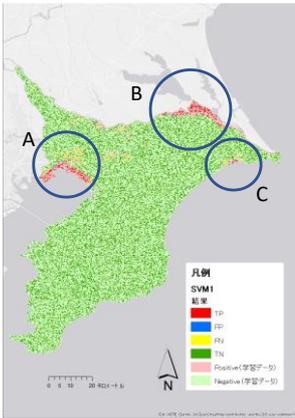


図-5 SVM1 の予測結果

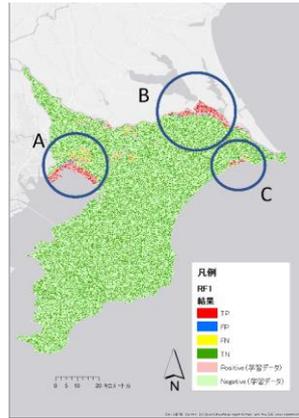


図-6 RF1 の予測結果

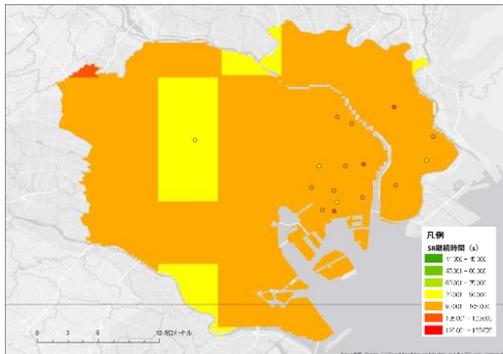


図-7 地震動継続時間の分布 (東京 23 区)

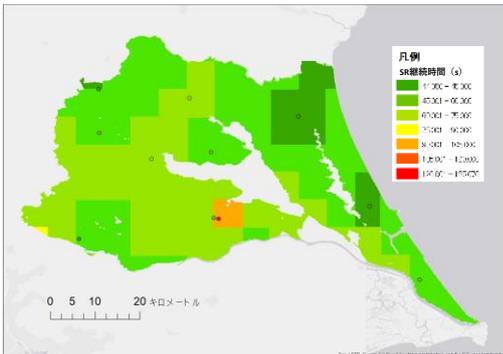


図-8 地震動継続時間の分布 (茨城南部)

たものが図-5、図-6である。どちらも東京湾岸と利根川周辺での液状化が集中している地域では的中率が高いものの、内陸部での見落としが多い。また、RF1は、SVM1と比べAの内陸部やCでの液状化を多く予測している。

4. 液状化発生予測手法の他地域への適用と予測結果の考察

前章の検討に基づき、最良としたRF1を東京23区と茨城南部に適応し、液状化予測結果の特徴を評価する。対象の地震は2011年東北地方太平洋沖地震の本震とし、使用データは千葉の際に用いたものと同一である。東京23区のメッシュ数は9981で、そのうち2%に当たる200メッシュで液状化が発生した。また、茨城南部のメッシュ数は34762で、そのうち8.5%に当たる2962メッシュで液状化が発生した。図-7、図-8に東京23区と茨城南部における地震動継続時間の分布を示す。

RF1を適用したところ、東京23区では正解率が0.942、再現率は0.520と比較的が高いが、適合率は0.178と低い。そ

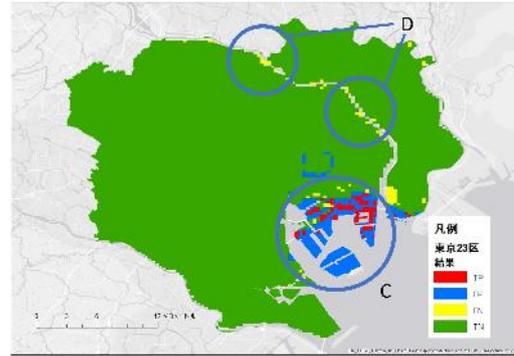


図-9 RF1による東京23区の予測結果

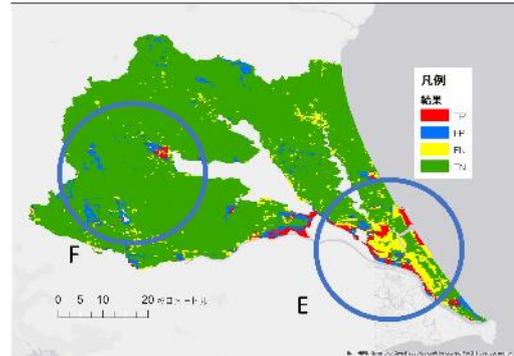


図-10 RF1による茨城南部の予測結果

のため、F値は0.226となった。図-9Cの東京湾側の埋立地で液状化を過剰に予測している。また、図-9Dの荒川沿いで液状化の見落としもみられた。

東京23区の埋立地では地震動継続時間が90~105秒であり、また計測震度は5.0~5.5である。この範囲は学習データとした千葉県の埋立地において液状化率が高いため、ほとんどで液状化ありと予測したと考えられる。

茨城南部の正答率は0.898である。図-10Eで液状化を見落としとしているメッシュ(FN)が多い。一方、図-10Fでは、液状化と誤って予測しているメッシュ(FP)がやや多い。全体として液状化が発生したメッシュのうち70%弱を見落とししたため、F値は0.336となった。

茨城南部の地震動継続時間は、30~75秒に分布している。このうち地震動継続時間が45~75秒のメッシュは、RF1の学習データである千葉県の液状化率が比較的高い。RF1の説明変数の影響度は地震動継続時間が最も大きいため、これらのメッシュは液状化と判別されやすい。このことから、液状化の誤抽出(FP)は、地震動継続時間が主原因と考えられる。

5. 他地域を加えたモデルの作成

前章の結果を踏まえると、学習データを特定の地域だけから取得するのではなく、広範囲からデータを取得し様々な液状化発生の特徴を含んだものとするのが重要と考えられる。そこで、学習データに東京23区、茨城南部のデータを加えたモデルRF4を作成した。学習データは、データ数をRF1と同程度とするため、千葉県の35%、東京23区の60%、茨城南部の60%の計54704メッシュとした。これを用いて分類器を作成し、各地域の残りのデータで適応性を検証する。

東京23区では、RF1と比べて図-11Aの埋立地での過剰な予測が抑えられ、正解率0.979、適合率0.370となった。しかし、液状化の見落としが増えたため再現率が0.132と

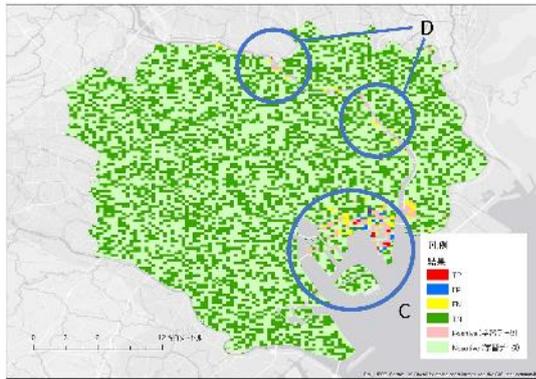


図-11 RF4による東京23区の予測結果

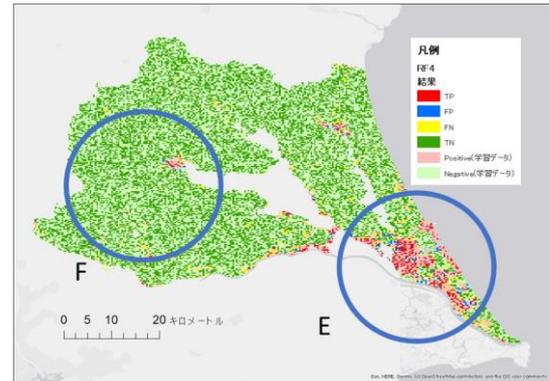


図-12 RF4による茨城南部の予測結果

低く、F値も0.194と低下した。

東京23区での液状化は埋立地での発生の割合が多い。東北地方太平洋沖地震では、竣工年が古い埋立地では液状化の発生が少ないことが指摘されている。東京23区では千葉県と比べこれに該当する埋立地が多く、これを判別する説明変数がないことが、精度が低下した原因と考えられる。また、その他の微地形では液状化率が低いため、精度の向上につながらなかったと考えられる。

茨城南側では正解率0.942、適合率0.722、再現率0.475、F値0.574と、RF1を適用した場合より全指標で精度が向上した。図-10と比較すると、図-12EでTPが増え、図-12FでFPが減少したことがわかる。

6. まとめ

本研究では、サポートベクターマシンとランダムフォレストを用いて、東北地方太平洋沖地震の際の千葉県を対象として、地震動継続時間を考慮した液状化発生予測モデルの構築を行った。SVMとRFのどちらにおいても、地震動継続時間を考慮することで精度が向上することがわかった。また、液状化の発生傾向を踏まえて微地形をグループ分けし、液状化発生予測モデルを構築した。しかし、液状化の見落としが増え、精度の向上にはつながらなかった。さらに、SVMとRFを比較すると、RFの方が高精度であった。

千葉県を学習データとして構築した液状化予測モデルを東京23区と茨城南側に適用し、予測結果の特徴を分析した。東京23区の液状化の再現率は0.520と比較的高かったが、東京湾側の埋立地のほぼ全域が液状化と予測された。このような予測結果が得られたのは、東京23区の埋立地における地震動継続時間と計測震度が、千葉県で液状化が集中的に発生した埋立地の地震動継続時間と計測震度の分布に完全に内包されることが理由と考えられる。また、茨城南側の西側では液状化を誤抽出し、東側では液状化を見逃した。西側の地域の地震動継続時間が、学習データである千葉県の液状化率が高い範囲と一致したため、液状化を過剰に抽出したと考えられる。

他地域を加えたモデルを作成したところ、茨城南側では精度が向上したものの、東京23区では低下した。東京23区では竣工年が古い埋立地が比較的多く、埋立地内でも液状化の有無が分かれたため、学習データに加えたものの精度が向上しなかったと考えられる。

このことから、本研究で構築した液状化発生予測モデルは、特定地域の特徴を強く反映したモデルとなった。千葉、茨城南側のテストデータに対する予測精度は、非常に高い。しかし、液状化発生の特徴が異なる東京23区への適用性

は低かった。汎用性の高い予測モデルを構築するには、学習データを増やすと同時に、地下水位や地盤改良など、液状化の発生に影響があると考えられる要素を考慮した説明変数を追加することが必要である。

参考文献

- 1) 若松和寿江, 先名重樹, 小澤京子: 2011年東北地方太平洋沖地震による液状化発生の特性, 日本地震工学会論文集, Vol.17, No. 1, pp. 43-62, 2017.
- 2) 猪股渉, 乗藤雄基, 石田栄介, 塚本博之, 山崎文雄: 東日本大震災における東京ガスの設備被害の概況と超高密度観測情報に基づく低圧ガス導管被害推定の精度検証, 日本地震工学会論文集, Vol. 13, No. 2, pp. 37-44, 2013.
- 3) 岩崎敏夫, 龍岡文夫, 常田賢一, 安田進: 地震時地盤駅状化の程度の予測について, 土と基礎, Vol. 28-4, No. 1164, pp. 23-29, 1980.
- 4) 松岡昌志, 若松加寿江, 橋本光史: 地形・地盤分類250mメッシュマップに基づく液状化危険度の推定手法, 日本地震工学会論文集, Vol. 11, No. 2, pp.20-39, 2011.
- 5) 若松加寿江, 松岡昌志: 全国統一基準による地形・地盤分類250mメッシュマップの構築とその利用, 日本地震工学会誌, No.18, pp. 35-38, 2013.
- 6) 桑原光平, 松岡昌志: 機械学習を用いた日本全国の液状化危険度の推定, 日本地震工学会論文集, Vol. 21, No. 2, pp. 270-289, 2021.
- 7) 柳瀬匡雄, 丸山喜久: 液状化地点の予測に向けた共分散構造分析とサポートベクターマシンの利用, 土木学会論文集 A1 (構造・地震工学), Vol. 75, No. 4, pp. I_133-I_143, 2019.
- 8) 若松加寿江, 先名重樹: 2011年東北地方太平洋沖地震による関東地方の液状化発生と土地条件, 日本地震工学会論文集, Vol.15, No.2, pp.25-44, 2015.
- 9) 先名重樹, 松岡昌志, 若松和寿江, 翠川三郎: 液状化発生率におよぼす強震動の継続時間と地域性の影響, 日本地震工学会論文集, Vol. 18, No. 2, pp. 82-94, 2018.
- 10) 能島暢呂, 高島拓也: 累積パワーに基づく地震動継続時間の等価振幅レベルに関する考察, 日本地震工学会論文集, Vol. 16, No. 6, pp. 1-15, 2016.
- 11) 竹内一郎, 烏山昌幸: サポートベクトルマシン (機械学習プロフェッショナルシリーズ), 講談社, 2015.
- 12) 波部斉: ランダムフォレスト, 情報処理学会研究報告, pp. 1-8, 2012. 平井有三: はじめてのパターン認識, 森北出版, 2012.