

機械学習に基づく液状化による千葉県全域の地盤沈下量予測

Prediction of liquefaction ground subsidence in Chiba Prefecture based on machine learning techniques

22WM1333 荻米 和樹
Kazuki Karimai

指導教員 丸山 喜久
劉 ウェン

SYNOPSIS

In this study, the authors developed a machine learning model to predict land subsidence caused by liquefaction resulting from an earthquake, with the aim of improving liquefaction hazard maps. A numerical model for predicting land subsidence due to liquefaction using XGBoost (eXtreme Gradient Boosting), which is one of the ensemble machine learning methods, was developed using the results estimated by the Nankai Trough Earthquake Model Examination Committee. We also visualized the features using SHAP (SHApely Additive exPlanations). Here, the numerical model was applied to estimate the land subsidence in Chiba Prefecture, where liquefaction was extensively observed after the 2011 Great East Japan Earthquake. The results showed a geotechnically valid map of expected land subsidence because significant subsidence was observed in the areas where liquefaction conditions were met.

1. 研究背景

液状化は、地震による被害の大きな原因であり、地盤の持つ支持力が急激に低下することで建物やインフラに壊滅的な被害をもたらす現象のことである。液状化に関する研究は多岐にわたり、土木工学、都市防災、災害リスク管理など、様々な分野で進められている。Seed and Idriss¹⁾は、液状化の発生条件として地盤の締固め度、土粒子の大きさ、地下水位などが重要であることを明らかにし、簡易的な液状化評価を提案した。2000 年以降、日本では実際の地震時のデータを基にした研究が活発となった。たとえば、2004 年の新潟県中越地震や 2011 年の東北地方太平洋沖地震、2016 年の熊本地震、2018 年の北海道胆振東部地震では、広範囲にわたり液状化被害が生じ、これらの事例を基にした多くの研究が行われている。

液状化現象による被害の 1 つとして地盤沈下が挙げられる。2011 年 3 月 11 日に発生した東北地方太平洋沖地震では、全国各地において大規模な液状化が発生し、路面の地盤沈下や噴砂現象による路面隆起によって緊急車両の通行を妨げたほか、避難速度の低下を招いた。液状化による被害が最も深刻だった千葉県浦安市では、道路の不同沈下や噴砂現象によって復旧作業が延滞し、全体的な地盤改良と補強工事が余儀なくされた。このように、液状化は避難行動についても大きな影響を与えることが懸念される。国土交通省は、全国各市町村の液状化危険度マップを取りまとめたポータルサイトを公開している。しかしながら、液状化危険度をカテゴリー変数として表現していることにとどまっており、液状化被害の程度を細かく捉えることが難しい。さらに、避難行動に影響のある地盤沈下量は定量的に評価されていない。

そこで本研究では、地震動データや地形データ、地盤情報などの既存データを統合的に用いて、液状化地盤沈下量を定量的に予測するモデルを構築する。同様に、説明変数の要因分析を行い、地盤沈下を引き起こす要因を明らかにすることを目的とする。

2. 使用データ

本研究では、南海トラフ巨大地震により想定される液状化地盤沈下量を目的変数に予測モデルを構築した。また、機械学習のうち教師あり学習を採用し、既存の教師データを予測モデルに学習させることで、目的変数を説明するための特徴量を抽出し、回帰問題を処理する。学習モデルのデータセットは、国土交通省²⁾が公開する 5 次メッシュ (250m メッシュ) 単位に整理した。教師データには、内閣府³⁾の南海トラフの巨大地震モデル検討会より検討された東海地震、東南海・南海地震の基本ケース、陸側ケース、東側ケース、西側ケースと、経験的手法に基づくケースの地盤沈下量を用いる (全データ 1326588 メッシュ)。また、メッシュ内の地震動については、南海トラフの巨大地震モデル検討会の地表震度 Ia および震度増分 dI から基盤震度 (Ia_min_dI) を計算し、説明変数に加える。さらに、国土数値情報²⁾から最高標高 (Max_Elev)、最低標高 (Min_Elev)、起伏量 (Relief)、最大傾斜角 (Max_Slop_Deg)、最小傾斜角 (Min_Slop_Deg) を、地盤の飽和度に関する特徴量として 5 次メッシュ中心点から海岸線までの最短距離 (Coast_Dist)、河川までの最短距離 (River_Dist) を説明変数として加えた。また、若松ら⁴⁾が整備したデータからは、深さ 30m までの地盤の平均 S 波速度 (AVS30)、微地形区分 (JCODE) を、山崎ら⁵⁾が整備したデータからは、水路からの高さを示すパラメータとして最寄り水路からの比高 (HAND) を追加した。地盤情報について、本研究では地盤の強度を示す指標である N 値を、国土情報検索サイト Kunijiban⁶⁾からダウンロードして使用する。具体的には地表面から 1m ごとに、深さ 10m までの N 値を説明変数として加える。ただし、N 値の取得深度が整数値とは限らないため、入力変数に合わせたデータ補間を行う。ここでは、データの収録状況を踏まえて、深さ 10m より深い地層の N 値が収録されていることを使用条件とする。データの補間方法について、まず地表面から N 値第一点、第二点境界までを直線で補間する。次に、第二点から第三点境界までを同様に直線補間、第 i 点から第 (i+1) 点境界までを繰返し補間し、補間線が深度 10m

に達するまで行う。この補間処理により得られる深度-N 値図について、対象メッシュから最も近いボーリング地点を、地表面から1m ずつ(N₁, N₂, N₃, N₄, N₅, N₆, N₇, N₈, N₉, N₁₀)に割り当てる。これらのデータは ArcGIS を用いて、5 次メッシュ単位に集約してから学習を行った。

多重共線性による特徴量の競合を防ぐために相関分析を行い、高い相関を持つ説明変数を除外した。各変数の相関は、ピアソンの相関係数(PCC)を分析することで推定できる⁷⁾。特徴の絶対値が PCC>0.95 であれば、予測モデルの精度に悪影響を及ぼす可能性がある。GIS データの相関における各説明変数のピアソン相関係数を示すヒートマップを図 1 に示す。図 1 から最高標高(Max_Elev)と最低標高(Min_Elev)、最小傾斜角(Min_Slop_Deg)と起伏量(Relief)との間に著しい正の相関が確認された。これに基づいて、本研究では、最低標高(Min_Elev)と起伏量(Relief)を除外した。

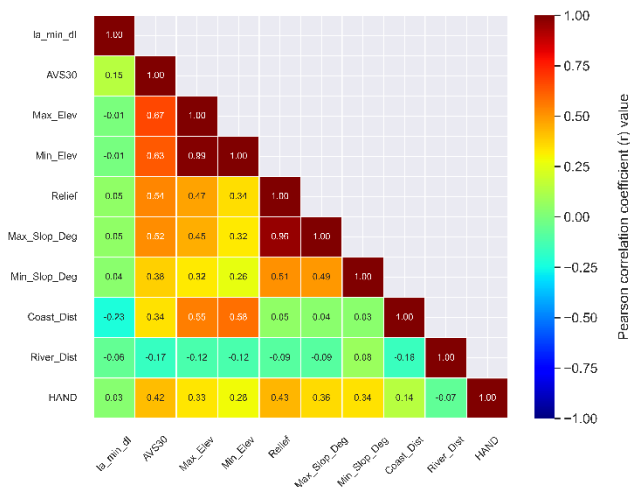


図 1 GIS データにおける各説明変数のピアソン相関係数

3. 予測モデルの構築

勾配ブースティング決定木は、勾配降下法(Gradient)、アンサンブル学習(Boosting)、決定木(Decision Tree)の手法を組み合わせた機械学習手法の 1 つであり、複数の弱学習器を組み合わせて単一の予測モデルを構築する。モデルの学習過程では、各段階で重みを繰り返し調整し、予測値と実際の値との誤差を減らすことで、説明変数と目的変数との関係を明示的に適合させる。このプロセスを経ることで、モデルは説明変数の特徴量抽出とデータの傾向パターンの学習を行い、新規のデータセットから予測・説明することが可能となる。特に勾配ブースティング決定木では、説明変数同士の積から能動的に交互作用特徴量を生成し、複数の変数との因果性を考慮することができる。このように、複数の特徴量が組み合わさった際の相互作用を高次元に表現できることから、良好な回帰予測モデルの構築が期待できる。本研究では、既存のデータを用いて、液状化による地盤沈下量を予測するモデルを構築した。ここでは、XGBoost (eXtreme Gradient Boosting)⁸⁾によるアルゴリズム手法を用いて予測を行う。XGBoost は、各ステップで勾配ベクトル関数に基づいて損失を減少させる。XGBoost は勾配ブースティング決定木に基づく Level-wise tree growth の学習手順を採用し、決定木を同時拡張させる。このアルゴリズム手法では、ステップごとに均等で決定木を構築するほか、欠損値の処理や過学習を防ぐための正則化など、多くの計算的特徴がある。

4. 予測モデルの精度検証

推論用データを用いて予測精度の検証を行った。ここでは式(1)により、絶対誤差を求める。

$$Absolute\ error = |y_{true} - y_{pred}| \quad (1)$$

正解値 y_{true} は各メッシュ属性データにおける推論用データの地盤沈下量を、予測値 y_{pred} は予測モデルに推論用データを入力した際の出力値を表している。液状化地盤沈下量の予測精度を図 2 に示す。図 2 は、推論用データを用いて検証した時の絶対誤差のヒストグラムおよび累積度数曲線である。これによると、液状化による地盤沈下量の予測値は、推論用データのうち 50%は ± 0.14 cm 以内の誤差に収まる。同様に、80%は ± 0.71 cm 以内、90%は ± 1.37 cm 以内の誤差に収まる。しかし、この結果は大局的な精度評価にすぎず、地盤沈下の大きさによっては予測精度に偏りがある可能性もある。そこで、予測値のバイアスを考慮した精度評価を行うこととする。

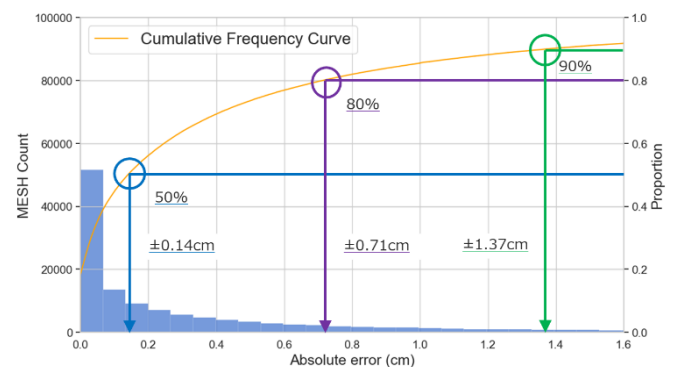


図 2 絶対誤差と累積度数曲線のパレート図

図 2 では推論用データ全体に対する絶対誤差を求めたが、ここでは絶対誤差を地盤沈下量の真値ごとに整理する(図 3)。地盤沈下あり(沈下量 > 0.0cm)のデータに関しては、分位数 10%ごとに分割した。図 3 は、真値を分位数ごとに分割したときの予測値の絶対誤差の箱ひげ図である。「箱」はデータの第 1 四分位数(Q1)から第 3 四分位数(Q3)までを表しており、図中のデータの 25%から 75%を集約する。図中の橙色線はデータの 50%である中央値を示す。箱から上下に延びる線は「ひげ」であり、データの範囲を示す。ひげの上限および下限は箱の長さの 1.5 倍以内ある最大値、最小値をそれぞれ示し、これを超える予測値は外れ値として扱われる。

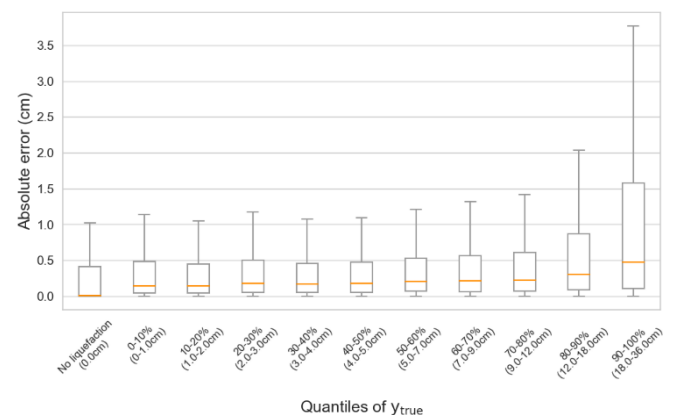


図 3 分位数ごとに分割させた予測値との絶対誤差

図3によると、地盤沈下量 0.0cm の絶対誤差は地盤沈下あり(沈下量 > 0.0cm)の場合よりも小さく、中央値はゼロである。地盤沈下量 0.0cm の方が教師データ数は多いため、モデルは変数間の特徴量をより多く抽出したと考えられる。地盤沈下量が 18.0cm より大きく 36.0cm 以下のメッシュは、推論用データの 75% で $\pm 1.6\text{cm}$ 以内の誤差に収まっている。分位数 0-80%(地盤沈下量が 0.0cm より大きく 12.0cm 未満)の絶対誤差は 0.6cm 以内で概ね一定である。しかし、分位数が 80%(地盤沈下量が 12.0cm より大きい)を超えると、絶対誤差は大きくなる。液状化による地盤沈下量が多い地点は、地盤の緩い箇所や細粒分の多い砂質土が影響しているものと想像されるが、本研究の説明変数には地盤の強さや土質構成等の情報が含まれていない。つまり図3の分位数 80-100%(地盤沈下量が 12.0cm より大きく 36.0cm 以下)のメッシュに関しては、本研究で用いた説明変数だけでは不十分であった可能性がある。液状化による地盤沈下量の推定には、地盤の力学的情報が重要であり、地質情報や地震動エネルギー、継続時間などの詳細データを説明変数に加えることが必要である。

5. 千葉県全域の液状化地盤沈下量マップ

気象庁⁹⁾の震度階級表に基づいて、各震度階級の液状化地盤沈下量を予測した。なお、予測モデルの地震動パラメータは量的変数の計測震度として学習させているため、震度階級を量的変数に置き換えてからモデルに入力する。震度6弱では $I=5.5$ 、震度5強では $I=5.0$ とした。千葉県全域の液状化地盤沈下量の予測に際して、液状化対象となるメッシュは千葉県が行う液状化危険度の検討手順¹⁰⁾を参考に抽出した。ここでは、山地、山麓地、丘陵、火山地、火山山麓地、火山性丘陵、岩石台地、砂礫質台地、ローム台地、水面に該当する微地形区分を非液状化メッシュとして扱う。ローム台地、水面は、若松ら⁴⁾の微地形区分の火山灰台地、沿岸海域として読み替えた。非液状化メッシュを除く液状化地盤沈下の予測マップを図4、図5に示す。

図4および図5によると、千葉県内房では東京湾沿岸に沿って地盤沈下が生じることを確認することができる。また、千葉県外房においても九十九里浜の陸側で広範囲に地盤沈下することが見られる。これらの地域での地盤は主に砂質土から成っており、特に埋立地区では液状化の生じやすい砂分が多く含まれるため、地盤沈下のリスクは高い。東日本大震災では浦安市や千葉市美浜区において噴砂や地盤沈下による建物の不同沈下が報告されているため、今後の巨大地震でも同地域において大規模な地盤沈下が予想される。

千葉県北部に着目すると、図5の基盤震度 $I=5.0$ でも地盤沈下 15.0cm 以上の激しい液状化が予想されている。この地域では一級河川の利根川が流れており、水域から近いメッシュでの地下水位は比較的浅く、弱い地震動でも間隙水圧の上昇は著しいと思われる。一般に、飽和地盤では地震波の伝播速度が低下することで地震波エネルギーが吸収されやすく、地盤増幅度は高い傾向にある。このような水域に近い地域での地盤沈下リスクは、これらの液状化要因の相互作用により非常に高いと考えられる。機械学習による予測モデルは説明変数間の特徴量を合成することが可能であるため、広範囲の地盤沈下量を推定するには合理的である。しかし、ピンポイントでの地盤沈下量の推定には予測精度上の観点から不向きである。現実的にはメッシュを細かくすることで程度対応可能であるが、データの拡充性や計算コストの問題から困難が予想される。そのため、正確な地盤沈下の推定には標準貫入試験やスクリーウェ

イト貫入試験(SWS 試験)の地盤調査を行う必要がある。本研究で作成した液状化ハザードマップは地震後の地盤沈下量を視覚的に示すものであり、避難経路の設計や地域社会のレジリエンス向上に期待できると思われる。

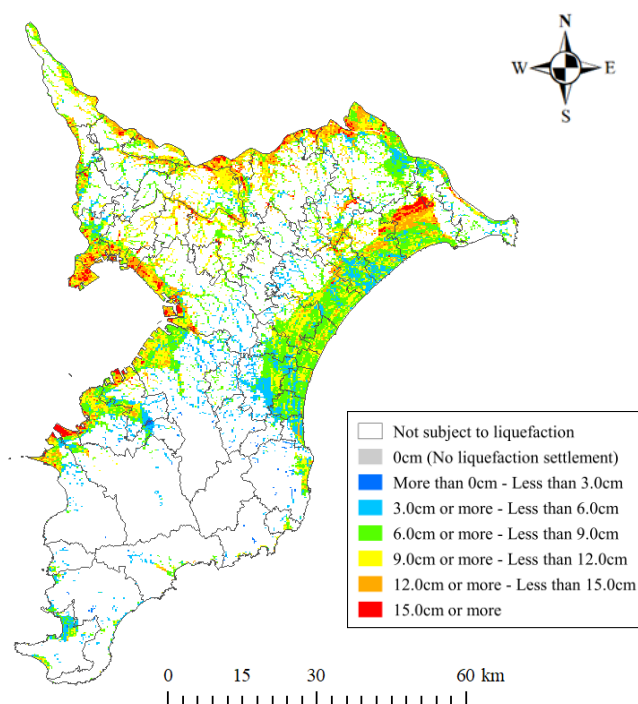


図4 基盤震度 $I=5.5$ の液状化地盤沈下量マップ

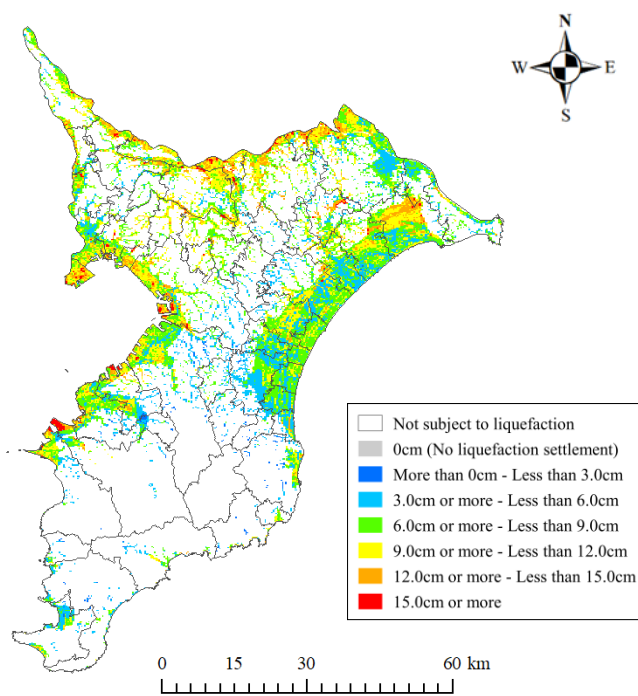


図5 基盤震度 $I=5.0$ の液状化地盤沈下量マップ

6. SHAP による予測結果の要因分析

SHapley Additive exPlanations (SHAP)¹¹⁾は、予測モデルの出力を解釈し理解するための可視化手法である。これは、協力ゲーム理論のシャープレイ値の概念に基づいている。本来は、連立のプレイヤー間の報酬を公平に分配するために定式化されたが、機械学習におけるシャープレイ値は、モデルの予測値に対する各特徴の寄与を可視化するために使用される。この手法によって、機械学習モデルの実用的な応用において解釈性が高まり、予測値に対する工学的説明が可能となる。これまで、SHAP による可視化手法は、機械学習モデルのブラックボックス問題を解消する手法として用いられている。SHAP の利点としては、異なる予測モデルでも特徴量が変化せず、モデルの構造を問わず一定であることが挙げられる。そのため、モデルの比較を主とする各説明変数の影響度を分析する際には、SHAP による解析手法が合理的だと思われる。

学習したモデルを用いて SHAP 値の計算を行った。地盤沈下量予測モデルにおける各説明変数の特徴量重要度を図 6 に示す。ここでは地盤沈下量に最も貢献する変数として上位 10 個を表示し、降順に示す。各説明変数の貢献度の大小関係は、地盤沈下量の大きさと関係がない。貢献度は、ある現象や結果に影響を与えた要因の程度を評価するための指標である。図 6 より、地盤沈下量に最も影響を与える変数は基盤震度(Ia_min_dI)であった。これは基盤震度の大きさによって地盤内の間隙水圧の上昇率が変化するために、地盤沈下量へのインパクトは大きいと考えられる。次に、深さ 30m までの地盤の平均 S 波速度(AVS30)、最高標高(Max_Elev)、10m 地点の N 値(N10)、海岸距離(Coast_Dist)の順に貢献度が大きく、地盤増幅度や標高、地盤飽和度に関する変数が上位であった。これらの変数は液状化の支配要因であるため、得られた結果は地盤工学的には妥当と考えられる。ボーリングデータの貢献度について、1 変数あたりの特徴量は少ないものの、図 6 の N 値(N10, N7, N8, N5, N9)グループにおける貢献度の総和は大きいと思われる。そのため、液状化による地盤沈下量の予測モデルにボーリングデータを説明変数として追加することに有効性はあると示唆される。地盤工学的に N 値は地盤の硬さの指標であるが、地形や地震動に関する特徴量と組み合わせることで、予測モデルの表現力向上に期待される。

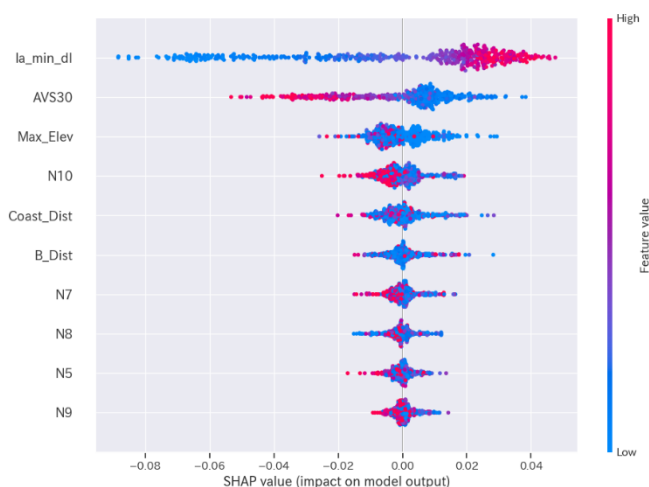


図 6 本研究で使用した説明変数の上位 10 変数における特徴量重要度(要約プロット)

7. 結論

本研究では、地震動データ、GIS データ、地盤データを用いて、千葉県全域の液状化地盤沈下量の予測マップを作成し、地盤沈下量に対する要因分析を行った。ここでは、XGBoost によるアルゴリズム手法を用いて機械学習モデルを構築し、特徴量抽出を行った。その結果、本研究で構築した予測モデルは全体的に高い精度を示し、教師データの正解値とモデルの予測値との適合性が良好であったことから、液状化地盤沈下量の予測には有効であることが確認された。また、SHAP の大局的説明より、工学的基盤における基盤震度と AVS30 の特徴量重要度が大きいことから、液状化地盤沈下量に対する影響は高いと示唆される。地盤の硬さを示す N 値が地盤沈下量へ大きく関与していることから、予測モデル構築の際にボーリングデータを説明変数として加える必要がある。GIS データや地形分類だけでは地盤情報を十分に表現できないために、ボーリング情報を取り入れることで、予測モデルの精度向上に貢献すると思われる。今後の課題は、地震動エネルギーや継続時間などの地震動パラメータの追加や、過去の地震記録から実際に地盤沈下が起こった地域での実地盤沈下量と比較検証を行い、予測モデルの精度向上に務める。

参考文献

- 1) Seed, H.B.; Idriss, I.M.: Simplified Procedure for Evaluating Soil Liquefaction Potential, Journal of the Soil Mechanics and Foundations Division, Vol.97, No.9, pp.1249-1273, 1971.
- 2) 国土交通省：国土数値情報ダウンロードサービス, <http://nlftp.mlit.go.jp/ksj/> (参照 2023-07-31).
- 3) 内閣府：南海トラフの巨大地震モデル検討会, <https://www.bousai.go.jp/jishin/nankai/model/> (参照 2023-07-31).
- 4) 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.
- 5) 山崎大, 富樫冨佳, 竹島滉, 佐山敬洋：日本全域高解像度の表面流向データ整備, 土木学会論文集 B1(水工学), Vol.74, No.5, pp.I_163-I_168, 2018.
- 6) 国土交通省, 国立研究開発法人土木研究所, 国立研究開発法人港湾空港技術研究所：国土地盤情報検索サイト, <http://www.kunijiban.pwri.go.jp> (参照 2023-12-22).
- 7) Lee, M.-H. Predicting and Analyzing the Fill Factor of Non-Fullerene Organic Solar Cells Based on Material Properties and Interpretable Machine-Learning Strategies. Solar Energy, Vol.267, p.112191, 2024.
- 8) Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), Association for Computing Machinery, pp. 785-794, 2016.
- 9) 気象庁：計測震度の算出方法, https://www.data.jma.go.jp/eqev/data/kyoshin/kaisetsu/calc_sindo.html (参照 2023-08-01).
- 10) 千葉県：液状化危険度の予測, <https://www.pref.chiba.lg.jp/bousai/jishin/higaichousa/documents/05syousu.pdf> (参照 2023-08-01).
- 11) Lundberg, S.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 22 May 2017; pp.4765-4774.