

Estimating the number of inland earthquakes using machine learning

*Tomoyuki Fujita¹, Hiroaki Takahashi², Ohzono Mako²

1. Graduate School of Science, Hokkaido University, 2. Institute of Seismology and Volcanology, Faculty of Science, Hokkaido University

内陸で発生する地震の長期評価として、地震本部が行っている主要活断層の地質調査等に基づくもの、GNSS観測による地殻ひずみ速度を用いたもの(Nishimura, 2022)、ETASモデルを用いたもの(Ogata, 2022)、地殻ひずみ速度と地震活動を用いたもの(Mazotti et al., 2011)などの開発が進んでいる。本研究では、内陸地震が複数のパラメータの分岐条件のもとに発生しているとの仮定に立ち、機械学習の手法として決定木を用いたアンサンブル学習であるランダムフォレストを用いて内陸地震の地震発生確率を推定することを検討した。日本列島を適当な大きさのグリッドに分割してグリッド内の各種特徴量(パラメータ)をリスト化した後、目的変数の推定を試みる。はじめは、特徴量のうち最大マグニチュード(M)のような適当なパラメータを目的変数とした教師あり学習を行って予測モデルの性能評価と改善を試行する。北海道を除く日本のグリッドを学習させたモデルを用いて、北海道のある一定期間の地震発生数の推定を行った。特徴量として、以下のパラメータを用いた；1) 各グリッドの活断層の有無；2) GR則のa,b値；3) 深さ10kmのP波、S波速度；4) 最大M；5) 面積ひずみ速度、最大せん断ひずみ速度。2)において、本研究では1997年以降の気象庁一元化地震カタログを利用する。注意することとして、地震活動には、前回の地震の余震活動が含まれている可能性があり、また群発的な活動を示す指標が必要となる。5)について、日本列島の地殻ひずみ速度は海溝での固着による弾性変形が卓越しているため空間的なハイパスフィルタを用いて短波長成分の利用を検討する。2)4)については震源位置の観測誤差の影響を考慮し、また地震数が少ない際のa,b値の推定が不安定になることを防ぐため周囲8グリッドを含む9グリッドのデータを用いて導出した。また、すべての特徴量において、特徴量が欠損しているグリッドについては周囲のグリッドのデータから最近傍法を用いて補間を行った。決定木の数、各木の深さ、特徴量の数などのハイパーパラメータについてはグリッドサーチにより決定した。さらに、結果についてどの特徴量が結果に大きく寄与したかを調査するため、ジニ不純度の減少量を求めその合計値から重要度を計算した。さらに、回帰モデルの精度評価として決定係数を用い評価した。本研究では、北海道全域にあたる226のグリッドにおいて、上記の特徴量を用いてそれぞれ地震発生数の推定を行った。結果として、十勝地域や道北地域の地震数が多いエリアでの推定数が大きくなっており、オホーツク海沿岸部や根室付近など地震数が少ないエリアについても推定値が小さくなる傾向がみられた。最も結果に寄与していたパラメータはGR則のa値、次点で最大Mであった。一方で、決定係数は0.19となり、用いた特徴量の傾向は推定結果に影響すると考えられるものの精度にはまだ問題があることが示された。今後は推定した地震数から支配的なパラメータを推定するとともに、マグニチュード5以上の地震の発生確率の推定を行う。また、評価手法について、推定の正誤判定の誤差の許容範囲の調整も行うことで、モデルの予測精度をより正しく評価できるよう検討する。