



| | |
|--------------|---|
| Title | Improved approaches to assess natural disaster susceptibility based on machine learning |
| Author(s) | Tang, Xianzhe |
| Citation | 大阪大学, 2021, 博士論文 |
| Version Type | |
| URL | https://hdl.handle.net/11094/85394 |
| rights | |
| Note | やむを得ない事由があると学位審査研究科が承認したため、全文に代えてその内容の要約を公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。 |

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

Abstract of Thesis

Name (TANG XIANZHE)

Title

Improved approaches to assess natural disaster susceptibility based on machine learning
(機械学習による自然災害感受性評価法の改良)

Abstract of Thesis

Natural hazards refer to natural process or phenomenon that may have negative impacts on society and natural disasters refer to social damages caused by natural hazards. The frequent occurrence of natural disasters brings considerable threats to people's live and global economy. It is critical to spatially assess the hazard susceptibility that contributes to reducing the related socio-economic loss.

Among the approaches used to spatially analyze natural hazards, the machine learning (ML)-based approach provides accurate evaluation results and avoids the data-scarcity at the same time. However, for the current ML-based assessments, there are existing several issues in both the technique and application. Firstly, only few studies consider the influence and efficient method for selecting discretization method (DM) for spatially continuous variables. Secondly, the existing method used for selecting negative samples for training classifiers cannot guarantee the classification performance and generalization capacity of the trained classifiers. Thirdly, a limited number of the employed positive samples negatively affects the quantity and quality of the selecting negative samples and the classification performance of the trained classifiers. Fourthly, most of the current ML-based natural hazard susceptibility assessments ignore to consider the hazard's temporal characteristics and consequence.

This thesis consisting of seven chapters is proposed to directly address the mentioned issues and thus improving the accuracy of the assessments based on ML.

Chapter 1 offers the needs of natural hazard susceptibility assessment and general approaches employed for such goals, reviews previous studies to clarify the current issue of the approaches especially based on ML, and states the objective and structure of this thesis.

Chapter 2 compares DMs suitable for discretize input variables and output susceptibility maps by taking a flood susceptibility assessment as a case. It proposes an index, Information Change Rate, that can quantify the suitability of DMs, contributing to select an optimal DM for each continuous variable and thus optimizing the susceptibility assessments.

Chapter 3 develops a negative sample selection method to obtain an optimal classifier, taking a wildfire susceptibility assessment as a case. It offers an optimized Repeatedly Random Undersampling that contributes to obtain a most accurate classifier with satisfied classification performance, leading to the improvement of the accuracy of susceptibility assessments.

Chapter 4 provides a positive sample increasing method to improve the ML-based flood susceptibility assessments. It presents an optimized seed sampling algorithm that can simulate potential inundated areas based on historical disaster locations and topographic features, contributing to increasing the number of positive samples with high-quality.

Chapter 5 demonstrates a new direction to broaden the ML-based natural hazard susceptibility assessments to damage assessment in an application domain through a case study of wildfire. It provides a framework that seasonally assesses the wildfire susceptibility and quantifies the caused economic losses to local ecosystems by types.

Chapter 6 discusses the implementations of the improved approaches presented in the previous chapters in the applications of ML-based hazard susceptibility assessments. It also argues the existing limitations and then offers the prospects for future studies.

Chapter 7 summarizes the achievements of the study and concludes the thesis.

Hence, based on these proposed solutions to address the existing issues, this thesis is anticipated to advance the view of consideration of ML-based natural disaster susceptibility assessments.

論文審査の結果の要旨及び担当者

| 氏 名 (T A N G X I A N Z H E) | | | |
|--|-----|----------------|---------|
| | (職) | 氏 名 | |
| 論文審査担当者 | 主 査 | 准教授 | 町村 尚 |
| | 副 査 | 教授 | 東海 明宏 |
| | 副 査 | 教授 | 近藤 明 |
| <p>論文審査の結果の要旨</p> <p>自然災害リスク管理において、災害ハザードの空間分布推定は効果的な災害対策の立案に不可欠である。単位期間中の災害ハザード発生確率は、災害感受性と定義されている。自然災害感受性の空間評価の手法には、物理的なプロセスに基づくモデル、統計モデル、機械学習によるモデルなどがあるが、その中で機械学習モデルは高い精度を持つ一方で、説明変数の離散化方法やトレーニングデータの選定方法に課題がある。また災害感受性マップを災害リスク管理に取り入れる方法論の検討も必要である。このような背景において、本論文は機械学習による自然災害感受性の評価精度を高めるアプローチを検討し、また災害感受性マップから災害による損失価値の期待値の時空間的評価に挑戦している。</p> <p>本論文は 7 章で構成される。第 1 章は自然災害感受性評価の必要性を示し、その様々な研究アプローチを既往研究よりレビューしている。またその中で機械学習を用いたアプローチの優位性と課題を述べ、本研究の目的と論文構成を提示している。</p> <p>第 2 章は機械学習による自然災害感受性評価において、説明変数の離散化法の選択手法を検討している。最適な離散化法選択の指標として Information Change Rate (ICR)を提案し、従来用いられていた Information Loss Rate (ILR)と比較している。中国 Wanan 県における氾濫災害感受性マップを事例として、ランダムフォレスト(RF)による分類器を構築しているが、ここで標高などの説明変数の 4 種の離散化法の中から ICR および ILR によって選択した方法を適用している。推定モデルの精度を AUC によって評価すると ILR よりも ICR が優れていることから、提案する指標が離散化法選択の課題を解決可能なことを示している。</p> <p>自然災害の空間的特性として、災害発生地点は一般に未発生地点より大幅に少ないため、自然災害感受性推定モデル構築のためのトレーニングデータの中では、ポジティブサンプルとネガティブサンプル候補地点の数の不平衡が大きい。第 3 章は、多数の候補地点から最適なモデルを得るためのネガティブサンプル選択法を検討している。AUC を目的関数としてランダムサンプリングを繰り返す Repeatedly Random Undersampling (RRU)法を提案し、従来用いられていた Single Random Undersampling (SRU)法と比較している。中国 Huichang 県における森林火災感受性マップを事例として、サポートベクターマシーン(SVM)と遺伝的アルゴリズム(GA)を組み合わせた GA-SVM 法による分類器を構築しているが、ここで分類器の最適化に SRU 法と RRU 法を適用している。AUC による最適分類器の精度評価によると SRU よりも RRU が優れていることから、提案する方法が最適分類器構築のためのネガティブサンプル選択の課題を解決可能なことを示している。</p> <p>第 4 章は、観測された災害発生地点に潜在的災害発生地点加えることでポジティブサンプルを補足するという、第 3 章と逆のアプローチを検討している。潜在的災害発生地点の抽出方法として、実際の災害発生点をシード点とし、災害感受性の主要な説明変数(氾濫災害に対しては標高)に基づいて 4 近傍または 8 近傍にシード点を拡大する Optimized</p> | | | |

Seed Spread Algorithm (OSSA)を提案している。中国 Guangzhou 市の 6 地域における氾濫災害感受性マップを事例として、RF、SVM、ロジスティック回帰分析による分類器を構築して OSSA の有無による AUC を比較している。3 つの機械学習法いずれでも OSSA によるポジティブサンプル増加によって AUC が高くなること、8 近傍より 4 近傍が優れていることがわかり、潜在的災害発生日点をポジティブサンプルに加えることで分類器の精度向上が可能なことを示している。

第 5 章は機械学習による災害感受性マップの応用として、森林火災による生態系サービス損失を 4 つの季節に分けて評価している。中国 Daxinganling を事例に、RRU 法を適用した機械学習による春夏秋冬それぞれの火災感受性マップと生態系クラスごとの各種生態系サービス価値をかけ合わせることで、生態系サービス価値の損失期待値を求めている。ここで生態系サービスは種別に純一次生産、降水量、土壌保全機能のいずれかの季節変化に応答するとし、サービス価値の季節性も考慮している。結果より、森林火災感受性に季節性があり秋季に最も高いこと、対象全体の生態系サービス価値は 824 億 USD であり夏季に最大であること、生態系サービス価値損失の期待値が 108 億 USD であること、常緑針葉樹の価値損失が最大であることを示し、災害感受性をリスク管理に利用するひとつの方法を示唆している。

第 6 章は第 2～5 章で示されている機械学習による災害感受性評価の課題の検討結果より、これらのアプローチの限界と将来の研究の方向性を論じている。

第 7 章は、本論文の達成度を検証し、結論としている。

以上のように、本論文は社会的要請が高い自然災害の感受性評価について、様々なアプローチによって機械学習による分類精度向上を達成するとともに、評価結果を定量的な被害期待値推計につなげている。

よって本論文は博士論文として価値あるものと認める。