論 文 の 内 容 の 要 旨

論文題目　Landslide susceptibility assessment using machine learning with emphasis on scaling and topographic representation issues
（空間スケールと地形の代表性を重視した機械学習による斜面崩壊の発生しやすさの評価）

氏　　名　ウッタム　ポウデル

Landslides are naturally occurring complex geological phenomena that cause significant damage in mountainous regions. Landslide susceptibility (LS) mapping is a requisite for safety against such sediment disasters caused by landslides, and considerable efforts have been exerted in this discipline. However, some key issues concerning the availability of thematic data, scaling of landslide causative factors, and the representation of landslide events still impose challenges. This research aimed to identify viable solutions to these issues.

Different parameters are generally used in LS studies (e.g., geology, soil depth, soil type, and land use) but limitations of data availability and resolution often restrict the replicability of such studies. Among the data required for LS studies, digital elevation models (DEMs) are currently the only dataset available globally at fine scales suitable for LS studies. This study examined the usefulness of a DEM-based LS analysis. One of the major challenges with this type of analysis is selecting an appropriate scale for LS studies due to the size heterogeneity and distribution of landslides, which requires identification of an optimal scale for landslide causative parameters. This study proposed a method to identify the optimum scale for each parameter and use multiple optimal parameter-scale combinations for LS mapping. Furthermore, the issue of topographic representation in a grid-based analysis arises because no raster cells within a landslide are equally responsible for landslide occurrence. Representation determines the sampling way of causative factors and thus affects further analysis. This study compared five different representation techniques: seed cells, center-cells, cells within the landslide boundary, cells within the depletion zone, and the dominant cells within the depletion zone (DCD), to identify the best representation technique. DCD is a new method of representation proposed in this study.

This study utilized Random Forest, a relatively new machine learning technique in the field of landslide research, together with 16 geomorphological parameters extracted from 10, 30, 60, 90, 120, 150, and 300 m DEMs and an inventory of historical landslides. The geomorphological parameters employed are those frequently used in existing landslide research. Two equal-sized (625 km^2) areas in Niigata and Ehime, Japan, with different geological and environmental settings and landslide density, were selected as study areas. The methodology was developed using high-resolution data available for Japan, and was successfully applied in a study area in Nepal where the quality of data is relatively low.

The usefulness of a DEM-based LS analysis was examined using two sets of models - with and without geological information. The results suggest that the addition of geological information leads only to a small increase in the prediction accuracy of the LS model in an area of high seismicity, and that the geological parameters are consistently ranked lower in importance than most other topographic parameters. Such an observation seems to reflect the coarser scale of the geological information used, and that the topography represented by a detailed DEM may include the effects of local geology. Accordingly, a DEM-based LS study is useful even if other high-quality datasets are unavailable, at least for rapidly eroding mountainous areas like those in Japan and Nepal.

A multi-resolution LS analysis technique was proposed to address the scaling issues. The method first determines the optimum scales for all parameters to best represent the conditions of slope failure. The parameters at different optimum scales are then brought together for the final LS mapping. The analysis of the scale and importance of the DEM-derived parameters revealed that while some parameters show similar importance and scale dependency for different regions, environmental differences result in variability between the study regions. The performance of LS models also suggests that the finest scale of analysis is not always the best. The proposed multi-resolution LS analysis permits higher accuracy LS mapping than any single-scale analysis.

The multi-resolution LS modeling was used together with five different representation techniques to evaluate the appropriateness of the techniques. The results indicate that the newly proposed DCD method always leads to higher performance in corresponding susceptibility models. This is logical because the dominant cell within the depletion zone probably represents the dominant process governing landslide initiation. The use of the proposed multi-scale approach together with the proposed DCD led to high prediction accuracies: 81.2% in Niigata and 83.27% in Ehime.

In summary, this dissertation work has suggested the usefulness of a DEM-based LS study with two newly developed methodologies for landslide modeling. The broad applicability of these findings should be examined in future research.