Earth Observation and Machine Learning for Geomorphological Mapping at Regional Scale

EARSeL 2019
Digital I Earth I Observation
Abstract
Corresponding Author:
eisank@grid-it.at

Clemens Eisank[1], Sebastian d'Oleire-Oltmanns[2], Volker Wichmann[3]

[1] GRID-IT, Austria
[2] Interfaculty Department of Geoinformatics - Z_GIS, University of Salzburg, Austria
[3] Laserdata, Austria

Keywords: Geomorphology, Landform mapping, Machine learning, Sentinel-2, TanDEM-X DEM, OBIA, Optimization

Abstract

Previous studies demonstrated the value of Earth Observation (EO) and machine learning for geomorphological and geological applications. Data acquired by recent satellite missions such as the Sentinels and TanDEM-X are well suited for extracting geomorphological features such as landforms and earth surface processes. Especially the following characteristics make these data useful for regional geomorphological mapping: global coverage, data consistency, spatial resolution of about 10 m, and provision free-of-charge. Machine learning provides a fast and efficient way to analyse the vast amount of satellite data. So far, in geomorphological mapping, machine learning approaches mainly focussed on extracting features in per pixel approaches.

We present an innovative approach for geomorphological mapping at regional scale through the application of machine learning within an object-based framework. The research was implemented in the project MorphoSAT (FFG ASAP, No. 859727).

The geodatabase included Sentinel-2 multispectral images, TanDEM-X DEM and a plethora of DEM derivatives such as slope and terrain wetness index. All datasets were collected/computed for a 100 km x 200 km study area in the Alps, and resampled to 10 m spatial resolution. Additionally, we gathered a large set of reference polygons for the training and validation of the selected geomorphological features. Selected features were glaciers, rock glaciers, landslides and debris cones. We split the reference data into a 25% training and 75% validation set. It is important to note that the reference represented a compound dataset that was created by merging data from four different institutional providers. Each provider used different data and techniques for digitizing the polygons. For the training polygons, we computed numerous polygon-specific image/morphometric data statistics as well as shape metrics. In total, we derived 324 variables. We employed several methods to reduce the variable space and to select the sets of variables with the most predictive power for the classification task. We identified ideal variable sets for multi-class mapping and for single-class extraction. The identified sets of training variables in combination with the training polygons were used to train five classifiers: random forest, classification and regression tree, support vector machines, k-nearest neighbours and naïve bayes. This resulted in a total of about 150 trained models.

Selected layers of the geodatabase were segmented using region-growing segmentation with different settings. Statistical optimization was performed to select three optimal segmentation scale levels out of

the many scales that have been produced. For the segments in these three levels, we computed the values of the variables required for prediction. The trained models were applied to the three segmentation levels to produce regional geomorphological classifications. In total, we generated about 200 polygon datasets with classified segments (Figure 1). Number- and area-based accuracy metrics, as well as visual analysis were the basis for assessing the quality of the geomorphological mapping results and the following degrees of freedom in the mapping system: classifier, number of training variables, size of training set, segmentation scale, single-class vs. multi-class mapping, software.

We are currently in the process of analysing the vast amount of derived metrics to evaluate and quantify the impact of the above-mentioned parameters. Eventually, we will be able to propose a set of parameters, which is ideal for regional geomorphological mapping based on machine learning and EO.Visual analysis of the results showed most promising results for glaciers followed by rock glaciers and debris cones. The most problematic class was landslides. Landslides were largely under-estimated or missed during the mapping. Obviously, the spectral and spatial data properties of landslide areas cannot be well distinguished from the data characteristics of other geomorphological features.

The results of the MorphoSAT project demonstrate that geomorphological features can be mapped at regional scale with acceptable quality, if variables with high predictive power can be found which in turn largely depends on the geodatabase. Since the proposed MorphoSAT mapping system relies on globally consistent EO data, the approach may be transferred to map geomorphological features in every region of the world. Moreover, global application is feasible with adequate processing infrastructure. The required EO data is already out there.

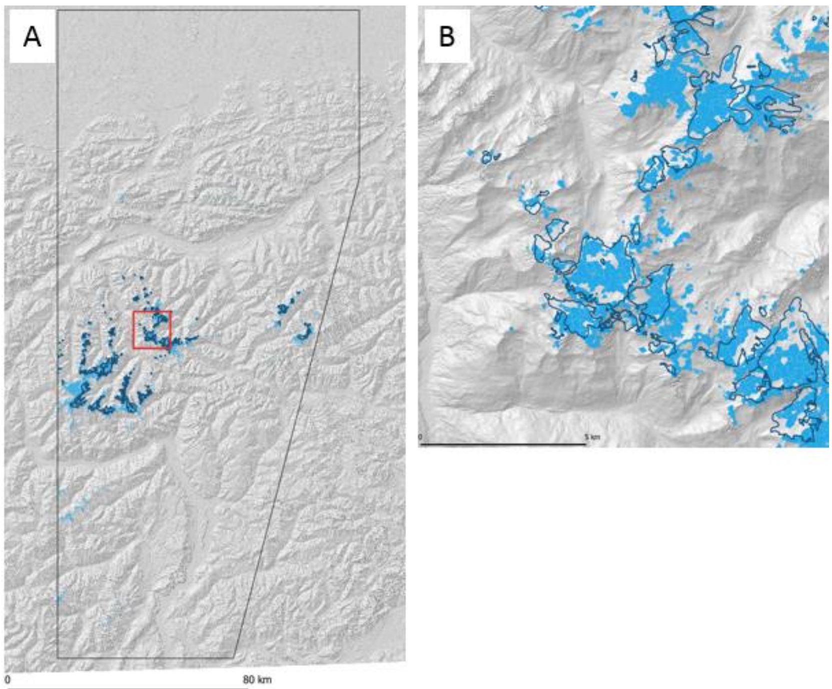

Figure 1. (A) Regional map of glaciers for the defined Alpine study area. Results produced by a random forest classifier that was trained with seven variables. Dark blue outlines show the boundaries of validation polygons. (B) Detailed presentation of results for the area inside the red rectangle in A.