Combining Participatory Mapping, Cloud Computing and Machine Learning for Mapping Climate Induced Landslide Susceptibility in Lembeh Island, North Sulawesi.

Safran Yusri, Endang Retnowati, Mikael Prastowo, Idris, & Fakhurrozi

Yayasan TERANGI
www.terangi.or.id
Lembeh Island in Bitung, North Sulawesi are mountainous and high slope area
- Landslide risk mapped by BNPB
- Local community experience landslide outside the predicted area
- Data difficult to find
Google Earth Engine

- Cloud platform for geospatial data analysis
- Includes terabytes of Earth Observation Data
- Ability to collect and manipulate large data
- Integrated Machine Learning
- Potential to be used for landslide modelling
Objectives

- Map landslide prone areas
- First test in implementation of landslide modelling with GEE and machine learning
- Provide input for disaster management in Lembeh Island
Methodology

- Lembeh Island in Bitung, North Sulawesi
- 3 villages: Kareko, Pintu Kota, and Pasir Panjang
- Field survey and interview to local community on landslide occurrences in the last 5 years
- Landslide modelling with machine learning (Pham 2016)
- Geodatabase includes Landslide Occurrence and Environmental variables
- Landslide occurrence were split to training (70%) and testing 30%) datasets
- Variables include: SRTM Digital Elevation Model, Sentinel 2 multispectral image, Climate Hazards Group InfraRed Precipitation with Station data
Methodology

- DEM -> Terrain Modelling -> Altitude, Slope, Curvature, Aspect
- Sentinel 2 -> Cloud Masking -> NDVI
- CHIRPS -> Data Reduction -> Kriging Interpolation -> 1 year Maximum and Mean
- Machine Learning Algorithm: CART (Breiman et al. 1984), Random Forest (Breiman 2001), GMO Maximum Entropy (Mann et al. 2009), Naive Bayes (Russel 1995), and SVM (Cortes & Vapnik 1995).
- Training Accuracy -> Confusion Matrix
- Testing Accuracy -> Error Matrix
Predicted landslide prone area is from 30.91 km² (Random Forest) to 40.25 km² (GMO Maxent).

With only 50 km² of land area, that means around 62% to 80% of land in Lembeh Island is susceptible to landslide.
Landslide Susceptibility Models
Landslide Susceptibility Models

| No | Algorithm                  | Training accuracy | Testing accuracy |
|----|----------------------------|-------------------|------------------|
| 1  | Random Forest              | 0.976             | 0.981            |
|    | Support Vector Machine     | 0.970             | 0.981            |
| 2  | CART                       | 1.000             | 0.981            |
| 3  | GMO Maxent                 | 0.964             | 0.981            |
| 4  | Naïve Bayes                | 0.893             | 0.904            |

Legend:
- Landslide Potential Area
Most landslide susceptible area are situated at the center of Lembeh Island, where the slope are 30° or higher.

Noorollahi (2018), slope is the parameter with the highest weight for determining landslide susceptibility.

Most of settlement are surrounded by high slope areas, thus, making these settlements prone to landslide impact.
Discussions

- High altitude area is more prone to landslide than lower altitude area.

- While most landslide reports collected are situated in the lower altitude due to the proximity to the settlements, the model successfully identifies landslide risk in high altitude areas.

- This shows that the models don’t experience overfitting.

[Map showing landslide and safe areas]
Variation in one year mean and maximum precipitation gave insignificant contribution toward classification.

This can be attributed to two factors, which are insufficient time scale and there was no significant difference of precipitation between areas.

Computation timed out error when calculating 30 years climate normal, thus needed a workaround.

GEE doesn’t support Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) directly.

Accuracy assessment only using Confusion Matrices.

Need to code the AUC and ROC by hand.
Conclusions

- Most area in Lembeh Island is prone to landslide, and the settlements are surrounded by it.
- Therefore, the residents must be aware, mitigate, and adapt to the hazard.
- GEE can be used to model landslides in areas where environmental data required for spatial analysis is not complete.
- Care must be taken due to evaluation of modelling accuracy can only use confusion matrices.
Thank you

Yayasan Terumbu Karang Indonesia (TERANGI)
Jalan Asyibaniah No. 105-106, RT. 03/RW.01, Pondok Jaya, Cipayung
Kota Depok, Jawa Barat 16438
Tel/Fax: (021) 29504088
e-mail: info@terangi.or.id
https://terangi.or.id
twitter: @terangi_