

## Identification of debris flow initiation zones using topographic model and airborne laser scanning data

### ABSTRACT

Empirical multivariate predictive models represent an important tool to estimate debris flow initiation areas. Most of the approaches used in modelling debris flows propagation and deposit phases required identifying release (starting point) area or source area. Initiation areas offer a good overview to point out where field investigation should be conducted to establish a detailed hazard map. These zones, usually, are arbitrarily chosen which affect the model outputs; hence, there is a need to have accurate and automated means of identifying the release area. In addition to this, the resolution of the terrain dataset also affects the results of the detection of source areas. In this study, airborne laser scanning (ALS) data was used because of its robustness in providing detailed terrain attributes at high resolution. Primary and secondary conditioning parameters were derived from digital elevation model (DEM) as input into the modelling process. Three models were executed at different spatial resolution scales: 5, 10 and 15 m, respectively. MARSpline multivariate data mining predictive approach was implemented using morphometric indices and topographical derived parameter as independent variables. A statistics validation was calculated to estimate the optimal pixel size, 1200 randomly sample data were generated from existing inventory data. Debris flows and no-debris flows were categorized, and the transform to continuous integer (1 and 0), respectively. To achieve this, the data set was divided into two, 70% (840) for the training dataset and 30% (360) for validation. The best model was selected based on the model performance using the generalized cross validation (GCV) and the receiver operating characteristic (ROC) curve/area under curve (AUC) values. Conditioning parameters were numerically optimized to identify the arbitrarily maximum model basis function for eleven variables, using MARSplines analysis (algorithm). The three most influencing topographic parameters identified are topographic roughness index (TRI), slope angle, and specific catchment area (SCA) with the percentage values of participation in the model of 100, 93, and 86%, respectively. The chosen function appeared to describe the analysed correlation sufficiently well. Consequently, three stages of optimization were made to determine the optimized source areas is possible with 10 m pixel size, 200 maximum basis functions and 3 maximum interactions, resulting into 82% ROC train and 80% test, GCV 0.189 and 85% correlation coefficient. The result will be of great contribution to the advancement of a broad understanding of the dynamics of debris flows hazard and mitigations at regional level which; that is resourceful for comprehensive slope management for safe urban planning decision-making process and debris flow disaster management.

**Keyword:** LiDAR; Debris flow; GIS; Remote sensing; Malaysia