# Exploratory Study of using Artificial Intelligence for Landslide Predictions

R.W.M. Cheung, H.W.M. Li*, E.K.H. Chu

Geotechnical Engineering Office, Civil Engineering and Development Department, Government of HKSAR, Hong Kong SAR, China

*Corresponding author
doi: https://doi.org/10.21467/proceedings.159.17

**ABSTRACT**

Riding on the comprehensive inventories of landslide-related data maintained by the Geotechnical Engineering Office (GEO) over the years, the GEO has initiated an exploratory study to enhance the existing landslide prediction models (i.e. Model A – landslide susceptibility model for natural terrain, and Model B – rainfall-landslide correlations for reported landslides on man-made slopes) with the application of machine learning (ML) and big data analytics. Model A adopted seven common ML algorithms to correlate the multitude of features (e.g. rainfall, geology, and some terrain-related features) with landslide in the natural terrain on the Lantau Island non-linearly. Domain knowledge of geotechnical and geological engineering was incorporated in the course of developing the ML model. The training and testing of the ML models used most of the available data as an approach to acquire realistic prediction of landslide probabilities out of an inherently acutely-imbalanced dataset. The applicability of some common evaluation metrics to this approach, and grid size effect were examined. Promising results with about three orders of magnitude enhancement to the model resolution were achieved. The use of ML on Model B is ongoing based on the knowledge and experience gained from Model A. This paper presents the latest progress of the exploratory study.

**Keywords:** Machine learning, Landslide susceptibility analysis, Rainfall-landslide correlation

## 1 Introduction

Landslide prediction models estimate the likelihood of landslide occurrence in an area based on a collection of landslide contributory factors. Over the years, the Geotechnical Engineering Office (GEO) has been continuously enhancing two territory-wide landslide prediction models which constitute the crucial parts of the slope safety system in Hong Kong (Cheung, 2021). The natural terrain landslide susceptibility model (Model A) predicts the spatial likelihood of the landslide occurrence of the natural terrain over the territory. It provides a rational and scientific basis to the formulation of the landslide risk management strategy including the spatial assessment of landslide risk and prioritisation of mitigation works, as well as land use planning. On the other hand, the rainfall-landslide correlation model for man-made slopes (Model B) provides real-time prediction of spatial landslide frequency and hence the total number of landslides over the territory in a rainstorm. It serves as the backbone of the Landslip Warning System (LWS) in Hong Kong, which was established in 1977 and is the first territory-wide early landslide warning system in the world (Chung *et al.*, 2022; Kong *et al.*, 2020). The LWS is intended to prompt the public to take precautionary measures to reduce their exposure to risk posed by landslide. It also supports the government's landslide emergency system. Both of the models were developed through data-driven analysis based on conventional statistical approach adopting predefined functions.

Recently, with the significant advancement of artificial intelligence (AI) which serves a highly capable yet readily accessible and cost-effective means for data driven analysis, the GEO has initiated an



Proceedings DOI: 10.21467/proceedings.159; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-965621-6-8

exploratory study of applying AI for landslide predictions. Focus is placed on enhancing the accuracy and resolution of the existing landslide prediction models with the use of machine learning (ML) technology coupled with big data analytics, which rides on the comprehensive inventories of territory-wide landslide-related data (e.g. landslide database, LiDAR data, rainfall data and geological maps) maintained by the GEO over the years.

## 2 Landslide Related Data Inventories

Good quality data with sufficient spatial and temporal resolutions are fundamental to data-driven analysis. In the past 40 years, the GEO has been collecting, and maintaining in light of technological advancement, a spectrum of landslide-related datasets in Hong Kong in a continuous manner. The geotechnical, geological and meteorological database pertinent to landslide prediction modelling that are maintained by the GEO are summarised in Table 1.

**Table 1:** *Summary of Geotechnical Database Pertinent to Landslide Prediction Modelling*

| Geotechnical Database | Available Information | Temporal and Spatial Resolution |
|---|---|---|
| LiDAR-based Digital Terrain Model (DTM) | Digital terrain models which can be converted into maps of various topographic information (e.g. gradient, curvature, aspect, catchment, etc.) | On average 4 data points/$m^2$ or 0.5m-grid for the territory-wide LiDAR survey undertaken in 2010 |
| Geological Maps | Geological data (e.g. solid geology, superficial geology and faults) based on a large amount of borehole logs, tunnel logs, cut slope exposures and geophysical survey records since 1980s. | Map scale = 1:20,000 |
| Rainfall | Spatial and temporal records of rainfall intensity of various durations | Average density of 1 rain gauge /10$km^2$<br>Readings taken at 1 min interval |
| Enhanced Natural Terrain Landslide Inventory (ENTLI) | Records of natural terrain landslides (including the unreported ones) identified from high-flight and low-flight aerial photographs (more than 100,000 landslide records) | Aerial photos temporal resolution = 1 year<br>ENTLI Database = update every 3 years |
| Reported Landslides | Records of landslides reported to the GEO since 1984 (about 10,000 records) | Annual review of reported landslides |
| Catalogue of Slopes | Details (e.g. location, geometry, geology, formation and failure history, maintenance responsibility, etc) of about 60,000 nos. registered man-made slopes | Continuous update when needed |

### 2.1 Digital Terrain Model (DTM)

The Airborne Light Detection and Ranging (LiDAR) technique can overcome the problem of views being obscured by dense vegetation, thereby obtaining the 'bare-earth' profile of the city and identifying geomorphological features more accurately. High resolution digital terrain model (DTM) as generated by multi-return LiDAR survey technique which can 'see through' vegetation in a territory-wide scale contains good quality topography-related information. The 0.5m-grid DTM was developed based on the 2010 territory-wide LiDAR survey. The DTM could be resampled into different grid sizes and converted into various topographic indices including gradient, plan and profile curvature, upslope catchment area and topographic position for landslide prediction modelling.

### 2.2 Geological Map

The GEO maintains a well-documented geological database of Hong Kong, which was developed based on extensive geological studies on a large amount of available borehole data (more than 300,000 ground

Proceedings DOI: 10.21467/proceedings.159
ISBN: 978-81-965621-6-8

Series: AIJR Proceedings
ISSN: 2582-3922

investigation stations over the territory), tunnel logs, cut slope exposures, geophysical survey records and other available geological information over the years. For the purpose of landslide prediction modelling, the lithology of the territory was categorised into three main groups (namely intrusive, volcanic, and sedimentary) based on the 1:20,000 solid and superficial geology maps so as to relate the engineering properties and thus landslide potential of the soil to their parent rocks. Other geological information e.g. distance to fault is also available but they were not adopted in the modelling further to assessment (see Section 3.2.3).

### 2.3 Rainfall Data

The landslides in Hong Kong are primarily induced by rainfall, which is highly correlated with the rainfall intensity and duration. In this regard, the GEO has been operating an extensive network of automatic rain gauge stations to collect real-time rainfall data at one-minute interval, which have been using for supporting the Landslip Warning System (Chung *et al.*, 2022; Kong *et al.*, 2020). Currently, the GEO Rain-gauge System consists of 90 automatic rain gauge stations in the network. Together with 31 automatic rain gauges managed by the Hong Kong Observatory (HKO) and the Drainage Services Department (DSD), the available real-time rainfall measurement at one-minute interval covers the whole territory of Hong Kong with an average density of one rain gauge per 10 $km^2$. With the data, spatial and temporal distributions of rainfall with different durations for a given rainstorm can be readily determined for subsequent analyses.

### 2.4 Enhanced Natural Terrain Landslide Inventory (ENTLI)

The ENTLI provides a comprehensive Geographic Information System (GIS)-based inventory of natural terrain landslides of the territory. It was firstly developed by identifying landslides from aerial photographs interpretation based on high-flight aerial photographs (taken at 2,400 m altitude or above) since 1945, and later supplemented with the results from the mapping of historical natural terrain landslides using low-flight aerial photographs (taken at lower than 2,400 m altitude) for an improved accuracy. Currently, over 100,000 landslides are recorded in the inventory. The ENTLI provides comprehensive and important data for landslide prediction modelling since landslides that occurred on natural hillsides and remote areas without affecting the public would have been missed otherwise, biasing the predictive model with the accessibility of the landslide locations. On the other hand, with its temporal resolution limited by the frequencies of aerial photo-taking and interpretation, the analysis of Model A adopting this database is bounded to be year-based as a result.

### 2.5 Reported Landslides

The GEO has been collecting landslide data through a reporting mechanism and conducting annual reviews of landslides and rainfalls in Hong Kong since 1984. On average, about 300 landslides are reported to the GEO every year. In general, most of the reported landslides are of a relatively small scale (i.e. less than 50 $m^3$ in volume), but some of them are sizeable (500 $m^3$ or more) (Ho & Cheung, 2021). To facilitate data management and manipulation, a computerised dataset of landslide records since 1984 has been established (Ho & Lau, 2008; Mak *et al.*, 2001). To date, about 10,000 landslides are recorded in the landslide database.

### 2.6 Catalogue of Slopes

All sizeable man-made slopes in Hong Kong (i.e. cut slopes, fill slopes and retaining walls) have been systemically identified from interpretation of historical aerial photographs and field inspections, and they are registered in a comprehensive Catalogue of Slopes (Ho & Cheung, 2021; Lam *et al.*, 1998).

The Catalogue of Slopes (the Catalogue) contains a wide range of useful information of the man-made slopes, such as geometry, geology, formation history, landslide records, maintenance responsibility and photographs, and the information contained in the Catalogue has been continuously updated. It provides the essential information of about 60,000 registered man-made slopes for landslide risk management.

## 3 Natural Terrain Landslide Susceptibility Model (Model A)

### 3.1 Existing landslide susceptibility model for the natural terrain of Hong Kong

A landslide susceptibility model predicts the spatial distribution of the landslide likelihood in an area with reference to its local terrain conditions (Brabb, 1984). The resulted susceptibility map enables the hazard zoning and the assessment the landslide risk of the area quantitatively, providing useful information for the decision makers to better understand and mitigate the landslide risks. The latest territory-wide landslide susceptibility model for the natural terrain of Hong Kong is given in Lo *et al.* (2022). This model discretised the natural terrain of Hong Kong into 5m x 5m grids (c.f. the total natural terrain area of Hong Kong is about 660 km$^2$), which were categorised into 144 classes and correlated to landslide susceptibility using conventional statistical approach. The classification of the grids was based on three landslide contributory factors (features), namely 1) eight classes of slope gradient, 2) three classes of lithology and 3) six classes of rainfall anomalies. A total of 24 years of landslide and rainfall data as recorded between year 1985 and 2008 were considered. Rainfall anomalies were characterised in terms of year-based normalised maximum rolling rainfall (NMRR) with rolling durations of 4-hour and 24-hour. NMRR is determined as the maximum rolling rainfall at a location normalised with the mean annual rainfall of the same location of a 30-year period from year 1977 to 2006. This model has an overall resolution of 4-5 orders of magnitude in terms of landslide density (no./km$^2$) (Figure 1). Based on this susceptibility model, a terrain-based landslide frequency map which predicts the annual theoretical number of landslides that would likely occur in a year for a given the mean annual frequency of rainfall occurrence was generated in Ko (2018).

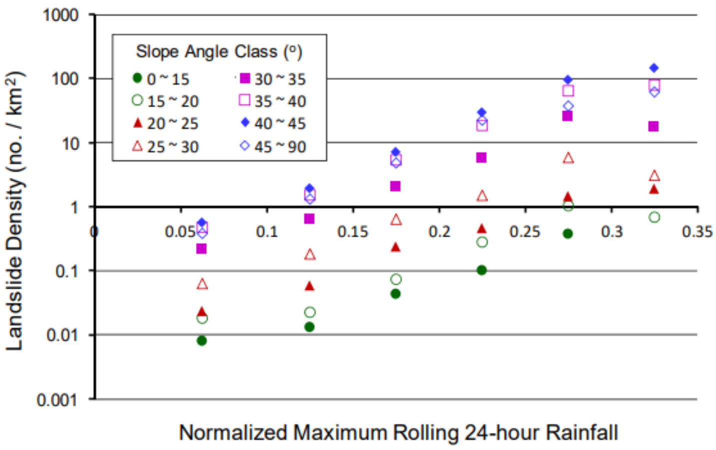



**Figure 1:** *Year-based Rainfall Landslide Correlation in* Lo *et al.* (2022) *for Combined Geology*

### 3.2 Machine learning-based natural terrain landslide susceptibility model

In order to explore the potential of applying ML in analysing the landslide susceptibility of natural terrain in Hong Kong, the GEO has undertaken a pilot study of Model A since 2021. This pilot study was carried out in two phases, with the first phase reported in Li *et al.* (2022). Further enhancement to the first phase ML model through reviewing the suitability of other algorithms and features was conducted in the second phase of the study. This paper reports the combined findings of the two phases of the study.

The study considered a pilot study area (Section 3.2.1) for the same period of year 1985-2008 as in Lo *et al.* (2022). The analysis was treated as grid-based binary classification problem (Section 3.3) with seven ML and deep learning algorithms considered (Section 3.2.2). A total of seven pertinent features were selected to train the ML model based on a tailor-made feature selection framework (Section 3.2.3). An optimal grid size of 5m x 5m was considered further to a review of the grid size effect (Section 3.2.4).

### 3.2.1 Pilot Study Area

The pilot study selected a 130 km$^2$ natural terrain on and around the Lantau Island (the pilot study area), which constitutes about one-fifth of the natural terrain area in Hong Kong (Figure 2). The pilot study area was selected for its high variability in topography-related and rainfall data, and the plentiful historical landslides available. Among the 24 years (i.e. 1985-2008) considered in the study, the study area experienced the most intense rainfall in 1993 and 2008, with the 24-hour maximum rolling rainfall of over 500 mm and 600 mm, respectively. The ENTLI recorded over 6,100 recent natural terrain landslides within this area.

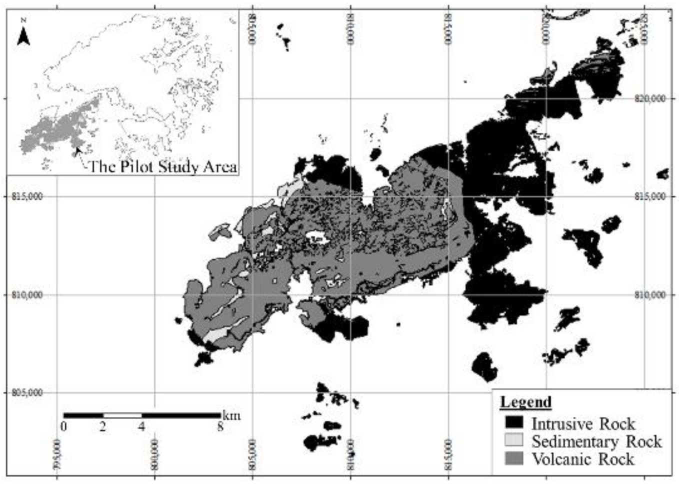



**Figure 2:** *The pilot Study Area*

### 3.2.2 Machine Learning Algorithms

The pilot study considered seven machine learning and deep learning algorithms which are commonly adopted for landslide susceptibility analysis. The key features of these algorithms are summarised in Table 2. In each of the study phases, the algorithms were applied to the same feature set to identify the best performing one(s). The assessment of the algorithms' performance was based on evaluation matrices as elaborated in Section 3.5.1. XGBoost was found to be the best performing algorithm in the Phase 1 study and it was tested again in Phase 2 to compare with the remaining algorithms.

**Table 2:** *Key Features of the Machine Learning/Deep Learning Algorithms Considered*

| Algorithm* | Phase studied | Characteristics | Remarks | Computational Efficiency** |
|---|---|---|---|---|
| DT | 1 | • Tree-based algorithm which partition data recursively to yield maximum information gain | • Less robust and is sensitive to the predictive data<br>• The patterns used to link the training datasets are highly interpretable | Very High |
| RF | 1 | • Tree-based algorithm<br>• Ensembles by the bagging method to lower variance and avoid model overfitting | • One of the best performing algorithms in ML competition before the introduction of XGBoost | Moderate |

| | | | | |
|---|---|---|---|---|
| XGB | 1,2 | • Tree-based algorithm<br>• Ensembles by boosting to improve robustness and generalizability<br>• Grows level-wise | • Introduced in 2016<br>• One of the best performing algorithms in ML competition | High |
| LGBM | 2 | • Tree-based algorithm<br>• Ensemble by boosting<br>• Grows leaf-wise to enhance computation efficiency | • Introduced in 2017<br>• Techniques to improve efficiency:<br>○ Gradient-based One-side Sampling (GOSS);<br>○ Exclusive Feature Bundling (EFB) | Moderate to High |
| ANN | 2 | • Deep learning algorithm<br>• Combines the activation functions in nodes through transformation with weight and bias applied | • Application not limited to structured data but also text, images, video and audio input | Moderate to High |
| SVM | 2 | • Works by maximizing the margin between decision boundary and data points | • Commonly adopted in ML-based LSA in literatures<br>• Determination of probability is highly resource demanding | Very Low |
| LR | 2 | • Model optimized based on maximum likelihood considering log(odds) | • Monotonic relationship between feature and probability assumed<br>• Cannot handle highly correlated features | Very High |

*DT = Decision Tree (Breiman *et al.*, 1984); RF = Random Forest (Breiman, 2001); XGB = XGBoost (Chen & Guestrin, 2016); LGBM = LightGBM (Ke *et al.*, 2017); ANN = Artificial Neural Network (Rosenblatt, 1958); SVM = Support Vector Machine (Cortes *et al.*, 1995); LR = Logistic Regression (Cox, 1958)

**the efficiency of computation was assessed specific to the dataset and approach taken in this pilot study

### 3.2.3 Feature Selection

While ML is often used in a way of maximizing the accuracy of predictions with the data mechanism behind unknown (Breiman, 2001), the GEO places much emphasis on ensuring the ML models remain physically meaningful for its prediction to be explainable without compromising prediction accuracy through critically incorporating domain knowledge to the model throughout the study. This is essential for model debugging and bias detection (Ma *et al.*, 2021), and the necessary understanding of the model based on which sensible decision can be made. In particular, a tailor-made framework was used to ensure potential features which are physically significant to landslide occurrences are selected for modelling only. This framework comprises the three criteria as follows:

1) the availability of quality dataset for the feature in terms of spatial and temporal coverage, resolution and accuracy;
2) the existence of a reasonable degree of statistical correlation between the feature and landslide occurrence; and
3) the above statistical correlation being explainable and consistent with the existing engineering principles (i.e. domain knowledge) on landslide occurrence.

Big data analytics techniques were adopted in compiling the suitable data for assessment against the framework, as well as for the ML model development. The characteristics of the features selected based on the above framework, and their physical significance to landslides are summarised in Table 3.

Proceedings DOI: 10.21467/proceedings.159
ISBN: 978-81-965621-6-8

Series: AIJR Proceedings
ISSN: 2582-3922

**Table 3:** *Characteristics of the Selected Features and Physical Significance to Landslides*

| Feature | Description | Physical Significance to Landsides |
|---|---|---|
| Slope Gradient | Inclination of slope | Affects the balance of stabilizing and destabilizing forces and thus the overall stability of a slope |
| Plan Curvature | The rate of change of slope gradient on horizontal plane | Influences the convergence and divergence of surface runoff and subsurface groundwater flow |
| Profile Curvature | The rate of change of slope gradient on vertical plane | Considered as a proxy to the break in slope that is assumed to be landslide related |
| Geology | Bedrock geology classified into three groups: Intrusive, Volcanic and Sedimentary | Related to the engineering properties of the soils derived from the parent rocks |
| Rainfall | Anomalies of rainfall characterised as normalised maximum rolling rainfall (NMRR) | Natural terrain landslides in Hong Kong were rainfall-induced; destabilising forces |
| Upslope Catchment Area | Contributory area of potential flow accumulating to the location | Indicates the degree of potential erosion of the area |
| Topographic Position Index | The difference in elevation of the grid and the average elevation of the surrounding ones | Indirectly reflects the terrain morphology and thus the susceptibility of landslide of the area |

Features commonly considered in literatures e.g. elevation, aspect, distance to drainage line, distance to fault, vegetation cover, etc were reviewed but eventually discarded for not fulfilling the selection framework.

### 3.3 Approach and workflow of modelling

The analysis was treated as grid-based binary classification problem in ML, which means the derived ML classifier predicts the landslide occurrence of a grid as a binary-dependent variables of positive or negative values (i.e. with and without landslide). This is a very commonly adopted approach among similar studies in Hong Kong (e.g. Ng *et al.*, 2021; Wang *et al.*, 2021). Since positive grids always constitute a very scarce portion of the data, the dataset for LSA is inherently acutely-imbalanced. Data sampling, while biasing the predicted probabilities of the classifier, is often used to achieve a more balanced dataset. However, in order to obtain a more realistic prediction of the landslide probability of the grids, this technique is not used in the pilot study. The alternative approach adopted enhances the overall resolution of the model and enables the prediction of the total number of landslides with simple mathematics.

The overall workflow of the modelling is illustrated in Figure 3. Geographical information system (GIS) software was used to prepare the feature datasets and they were then converted from raster to structured format for modelling, and vice versa for visualisation of the model output. All ML processing tasks including data pre-processing, model training and model performance evaluation were carried out with python coding. These tasks were carried out on a cloud platform given the enormous amount of data to be handled. The resampling of data for model training, validation and testing is also shown in the same figure.

The ML models are trained with the input dataset comprising all the available data except those that were reserved for testing. The hyperparameters of the ML models were tuned using five-fold cross validation (CV), with the training data randomly split to 7:2 to serve as the training and validation data in each fold. Stratified sampling based on landslide occurrence (i.e. landsliding or non-landsliding) was adopted such that same ratio of landsliding to non-landsliding data was maintained.

As for the model performance evaluation, this pilot study adopted an additional set of testing data (TD1) for evaluating model performance as compared with literatures. It comprises the data from Year 1993 and 2007 data which correspond to a high and a moderate rainfall scenarios, respectively. This testing data TD1 mimics future data that were unseen by the models to test their abilities in making forward prediction. On the other hand, testing dataset TD2 randomly extracted from the remaining 22 years of data was of nature similar to those considered in literatures. It is adopted to enable meaningful comparison and benchmarking of the model performance with literatures.

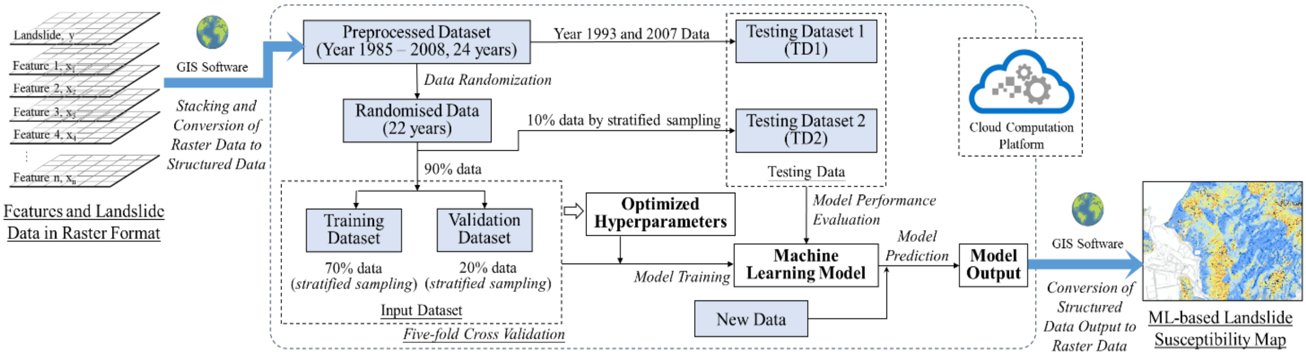



**Figure 3:** *Overall Workflow and Resampling of Data of the Pilot Study*

## 3.4 Grid size effect

An optimal grid size improves the predictive power of the ML model through effectively capturing the morphological conditions affecting the stability of the slopes (Martinello *et al.*, 2023). In order to study the appropriateness of the adopted grid size of 5m in this pilot study, the performance of a 5m-grid XGB model (the best performing algorithm in Stage 1 study) is compared with that of a 10m-grid XGB model adopting the same feature set in Table 4. Grid size of 10 m was considered since it covers about 80% of the landslide scars according to ENTLI. ROC AUC to be discussed in Section 3.5.1 is used to measure the performance of the model. It is evident from Table 4 that 5m-grid gives a better model performance and hence it is considered more effective in capturing the landslide-related morphological conditions.

**Table 4:** *ROC AUC of the XGB Models based on Different Grid Sizes (5m vs 10m)*

| Grid Size | Testing Data 1 (TD1) | Testing Data 2 (TD2) |
|---|---|---|
| 5 m x 5 m | 0.915 | 0.973 |
| 10 m x 10 m | 0.896 | 0.961 |

## 3.5 Results and discussion

The performance of the ML models was assessed in respect of (1) classification accuracy, (2) accuracy of the predicted landslide probability and (3) the degree of enhancement to model resolution.

### 3.5.1 Classification accuracy

The performance of the ML models in respect of classification accuracy is measured using the Area Under Curve (AUC) of the Receiver Operating Characteristic Curve (ROC), which is the most commonly used index for assessing the performance of a ML model (Spackman, 1989). Table 5 summarizes the results of the seven ML algorithms being tested in this pilot study. The predicted class of a grid is determined based on its predicted probability of the positive class with reference to a classification threshold. An ROC curve plots the true positive rate (TPR, i.e. the proportion of the truly classified actual positive data) against the false positive rate (FPR, i.e. the proportion of falsely classified actual negative data) for the full range of classification threshold from 0 to 1 (Figure 4). The

Proceedings DOI: 10.21467/proceedings.159
ISBN: 978-81-965621-6-8

Series: AIJR Proceedings
ISSN: 2582-3922

concept of TPR and FPR for a given threshold is illustrated in a confusion matrix in the same figure. A higher ROC AUC value indicates a better model performance in classification accuracy.

**Table 5:** *Summary of the ROC AUC Values Achieved by ML Models Adopting Different Algorithms*

| Data* | DT | RF | XGB | LGBM | ANN | LR | SVM |
|---|---|---|---|---|---|---|---|
| TD1 | 0.876 | 0.908 | 0.915 | 0.910 | 0.913 | 0.865 | 0.796 |
| TD2 | 0.944 | 0.967 | 0.973 | 0.973 | 0.964 | 0.953 | 0.829 |

* see Section 3.3 for the discussion on the characteristics of the testing datasets TD1 and TD2

The overall ROC AUC achieved by the models is comparable to the range of values reported in the literature (see Figure 4). LR and SVM had the lowest ROC AUC amongst the studied algorithms. The relatively low ROC AUC of LR is probably attributable to its intrinsic property which correlates data in a monotonic manner, and thus limits the model's predictive capability. On the other hand, as compared with the other algorithms, SVM required significantly longer running time under the adopted approach which makes it the least competitive model. The reliability of predicted landslide probability by models achieving the highest classification accuracy, including XGB, LGBM and ANN, were further assessed in the next section.

**Predictive Class**

| Actual Class | Positive | Negative |
|---|---|---|
| Positive | TP (True Positive) | FN (False Negative) |
| Negative | FP (False Positive) | TN (True Negative) |

Range of ROC AUC achieved in literatures

| Study | ROC AUC* |
|---|---|
| Liu et al. (2019) | 0.87 – 0.99 |
| Ng et al. (2021) | 0.872 – 0.901 |
| Wang et al. (2021) | 0.949 – 0.972 |
| Xiao et al. (2022) | 0.93 – 0.94 |

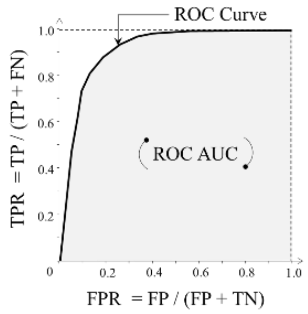



**Figure 4:** *Confusion Matrix and Receiver Operating Characteristic (ROC) Curve*

Other available evaluation metrics for the evaluation of classification accuracy includes accuracy, sensitivity, specificity and precision, etc. Most of them combines the distribution of these true positive (TP), false positive (FP), true negative (TN) and false negative (FN) data in various manner. A threshold of 0.5 is often considered for models trained with sampled data. Nonetheless, the present review indicates that the use of these metrics which are based on a single classification threshold for assessing model performance can be misleading for models derived from an imbalanced dataset and therefore they were not used in this study. On the other hand, accuracy of the predicted probability of the ML models were examined in the next section in lieu of using metrics e.g. brier's score or log-loss for a more detailed assessment.

### 3.5.2 Landslide probability prediction

The approach adopted in this study is considered to enable a more realistic prediction of the landslide probability, which can be directly taken as the probability of the positive class of individual grids. This is validated by comparing the values of the probability of the positive class as predicted by the model and the actual landslide probability of the grids. Figure 5 compares the two values of the XGB, LGBM and ANN models accordingly, with a linear relationship with a gradient of unity indicating good agreement among the two quantities. It is noted that XGB and ANN models give quite a reasonable prediction of landslide probability for values higher than $5 \times 10^{-6}$, while LGBM model basically under predicts in all cases. Neither XGB nor LGBM models are able to differentiate data points with predicted probability of landslide beyond 5x10$^{-6}$. While ANN model gives landslide probability predictions beyond 5x10$^{-6}$, those are under predictions. Nonetheless, differentiating areas with P(LS) lower than 5x10$^{-6}$ is seldom pursued in applications.

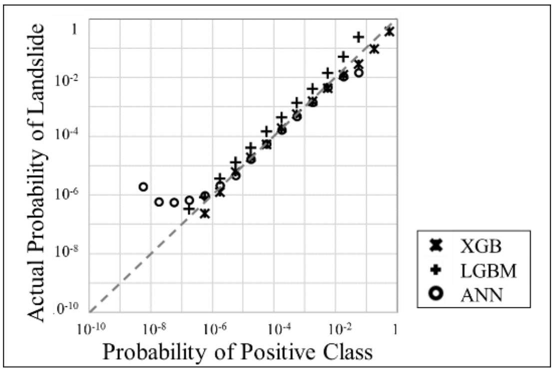
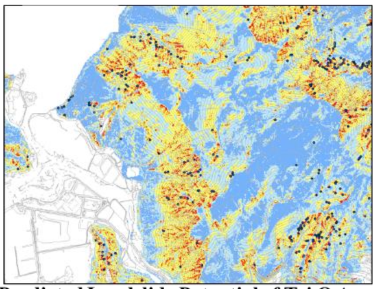



**Figure 5:** *Predicted Probability of Positive Class vs Actual Landslide Probability of the XGB, LGBM and ANN Models*

### 3.5.3 Degree of enhancement to model resolution

The resolution of the susceptibility model is dependent on the relative landslide susceptibility among grids, while the susceptibility of a grid is determined with reference to the predicted probability of landslide in this study. With the inclusion of multitude of features to the model enabled by the use of ML, this section looks at the enhancement to the resolution of the model as a result through assessing the range of the landslide probability of the entire pilot study area as predicted by the ML models. XGB model is considered since it provides the most reliable landslide probability predictions according to Section 3.5.2. The resolution of the models accounting for seven selected features in five different classes of the 24-hour NMRR is compared to that of a reference model in Figure 6. This reference model is an XGB model considering the three features adopted in the Lo *et al.* (2022) study (i.e. gradient, lithology and rainfall anomalies) for benchmarking purpose. The resolution of the model is found to be enhanced by two to three orders of magnitude, attributing to the introduction of four additional terrain-related features. The resulted spatial distribution of the predicted landslide probability in the Tai O area subject to 24h-NMRR of 0.275 produced by the 2022 XGB model is given in Figure 6 for visual illustration.



**Predicted Landslide Potential of Tai O Area**

**Range of Predicted Landslide Probability of the Pilot Study Area**

| ML Model | | RF Class I (0.025 – 0.10)* | RF Class II (0.10 – 0.15) | RF Class III (0.15 – 0.20) | RF Class IV (0.20 – 0.25) | RF Class V (0.25 – 0.30) |
|---|---|---|---|---|---|---|
| Reference Model | Min | 2.12 x 10$^{-7}$ | 2.54 x 10$^{-7}$ | 6.35 x 10$^{-7}$ | 3.45 x 10$^{-6}$ | 9.61 x 10$^{-6}$ |
| | Max | 2.04 x 10$^{-5}$ | 3.36 x 10$^{-5}$ | 2.10 x 10$^{-4}$ | 2.68 x 10$^{-3}$ | 2.56 x 10$^{-3}$ |
| 2022 XGB Model | Min | 2.80 x 10$^{-7}$ | 2.97 x 10$^{-7}$ | 4.64 x 10$^{-7}$ | 7.89 x 10$^{-7}$ | 9.02 x 10$^{-7}$ |
| | Max | 2.74 x 10$^{-3}$ | 2.20 x 10$^{-3}$ | 7.76 x 10$^{-3}$ | 3.70 x 10$^{-2}$ | 5.84 x 10$^{-1}$ |

* *values in bracket denote the range of 24-hour NMRR represented by the rainfall (RF) class*

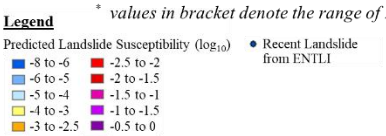



**Figure 6:** *Range of Predicted Landslide Probability of the Pilot Study Area and an Extract of Predicted Landslide Potential of the Tai O Area*

# 4 Correlation Model for Rainfall and Reported Landslides on Man-Made Slopes (Model B)

## 4.1 Landslide prediction for Landslip Warning System

As one of landslide risk management tools in the Slope Safety System in Hong Kong, a territory-wide Landslip Warning System has been established and implemented, with an aim to alerting the public of

Proceedings DOI: 10.21467/proceedings.159
ISBN: 978-81-965621-6-8

Series: AIJR Proceedings
ISSN: 2582-3922

possible landslide risk due to heavy rainstorms and supporting the landslide emergency system within the government to deal with landslide incidents. Landslide prediction models have been developed to the operation of the Landslip Warning System during heavy rainstorms. A Landslip Warning may be issued when the prediction of landslides reaches the threshold level.

It is recognised that a significant number of landslides at remote areas, such as natural hillsides at country parks, may not be found and reported, so the number of reported landslides in a rainstorm may only reflect a portion of the full figure of landslides occurred. However, as these landslides at remote areas may not have significant impacts on public safety, the prediction models for the Landslip Warning System thus only focus on predicting the number of reported landslides in a rainstorm, instead of covering all landslides occurred.

Rainfall is the major trigger of landslides in Hong Kong. Rainfall severity of a given rainstorm could be characterised in different ways, such as maximum rolling rainfall and antecedent rainfall with different durations, and considered into a landslide prediction model. As different types of man-made slopes may perform differently under the same rainfall conditions, characteristics of the man-made slopes, such as slope geometry, slope forming materials and level of geotechnical input, are potential contributory factors of landslide occurrence and are commonly incorporated into the prediction model.

### 4.2 Current prediction models for reported landslides on man-made slopes

Four generations of landslide prediction models have been developed for the GEO Landslip Warning System since 1977 (Chung *et al.*, 2022; Kong *et al.*, 2020). Along with the development of the GEO Raingauge System and the establishment of the Catalogue of Slopes since the 1980s, the prediction models have been continuously reviewed and updated based on the available datasets of landslides, rainfall and man-made slopes, to reflect the changing of landslide risk through urban development in Hong Kong. In the latest model, spatial distributions of registered man-made slopes and reported landslides, as well as spatial and temporal variations of rainfall intensities were adopted to develop bi-linear correlation models between landslide frequency and maximum rolling 24-hour rainfall for four types of man-made slopes (i.e. soil cut, rock cut, fill slope and retaining wall) to predict the number of landslides. The latest correlation models were updated in 2012 based on the data for the period of 1996 to 2010, followed by another review in 2018 with no further update to the correlations using the data up to 2016. Figure 7 shows the current correlation models between landslide frequency and maximum rolling 24-hour rainfall for four types of man-made slopes.

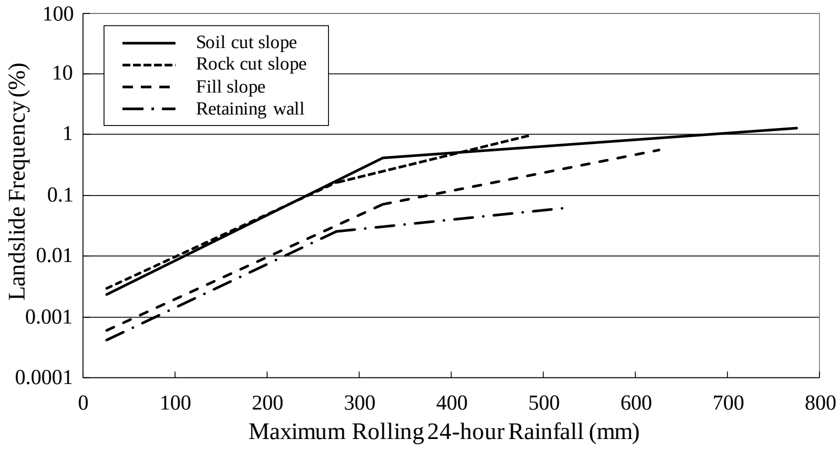



**Figure 7:** *Correlations between landslide frequency and max. rolling 24-hour rainfall for four types of man-made slopes*

### 4.3 Opportunities of machine learning

The GEO has launched an exploratory study to apply ML to revisit the relationship between rainfall and reported landslides on man-made slopes and seek opportunities to enhance the current capability in landslide prediction. Alternative ways of characterising rainfall severity and additional landslide contributory factors will be explored using ML in this study, with a view to improving the performance of the prediction models.

The maximum rolling 24-hour rainfall was adopted in the current landslide prediction models, as being a parameter strongly correlating failures of man-made slopes (Yu, 2004; Pun *et al.*, 1999) as well as a convenient parameter to characterise the intensity of rainfall in a given rainstorm. It was noted that maximum rolling 24-hour rainfall might not be a good indicator of landslide potential in some occasions (Yu, 2004). To gain more insight, it is worth exploring ways of characterising rainfall severity in correlating the landslide frequency of man-made slopes in a given rainstorm using ML tools. Maximum rolling rainfall with durations other than 24-hour and antecedent rainfall with different durations are examples to be explored whether or not they have statistical links with the landslide frequency in a given rainstorm.

Furthermore, additional potential landslide contributory factors that have not been considered before could be introduced to enhance the capability in landslide prediction model using the ML tools, which enable non-linear and multivariate analyses. The high-quality inventories of territory-wide landslide-related data established and maintained since the 1980s could be utilised for identifying potential landslide contributory factors. Apart from slope types and slope forming materials considered in the current prediction model, incorporating other slope information, such as slope angle and level of geotechnical input, into a ML model may give higher prediction accuracy. Although a wide range of data in various inventories are available, particular attention should be paid in selecting suitable features with due considerations of their physical meaning, reliability and accuracy. Careful feature engineering process should be carried out to make sure that selected features are consistent with our domain knowledge on landslide occurrence.

### 4.4 On-going development work

The work plan of the exploratory work of applying ML in the prediction of landslides on man-made slopes has been formulated. A slope-based binary classification analysis is conducted to formulate a ML model of a storm-based prediction of the reported landslides on the man-made slopes. Similar approach as in Model A is adopted to handle the highly imbalanced data of the reported landslides under a rainstorm event.

To develop the ML-based landslide prediction model for man-made slopes, the datasets of reported landslides, rainfall and man-made slopes have been compiled, extracted and pre-processed, and are ready for model construction. An assessment for feature selection is being conducted to select pertinent features with high quality and statistical and physical relevance for inclusion in the ML model. As a first attempt, a ML model incorporating the features considered in the current bi-linear correlation models for landslide prediction (i.e. maximum rolling 24-hour rainfall, slope types and slope forming materials) is being developed using some tree-based ensemble ML algorithms, such as Random Forest and XGBoost. The dataset resampling for model training and testing is presented in Figure 8. The datasets of about 200 rainstorm events from 1996 to 2010 are extracted and pre-processed. The dataset is resampled into a training dataset, a validation dataset, and two testing datasets, which comprise the data of 10% of rainstorm events (i.e. Testing Dataset 1), and 10% of the data randomly selected from

Proceedings DOI: 10.21467/proceedings.159
ISBN: 978-81-965621-6-8

Series: AIJR Proceedings
ISSN: 2582-3922

the remaining rainstorm events in a stratified manner (i.e. Testing Dataset 2), respectively. Currently, model training using the ML algorithms with cross validation and optimising their hyper-parameters are in progress. The performance of the trained ML models will be evaluated using model metrics, such as ROC AUC and Brier Score to compare the accuracy of the ML algorithms using Testing Dataset 2, and performance metrics, such as $R^2$ and Root Mean Squared Error to compare predicted numbers of reported landslides using the ML models and bi-linear correlation models with actual numbers of reported landslides using Testing Dataset 1.

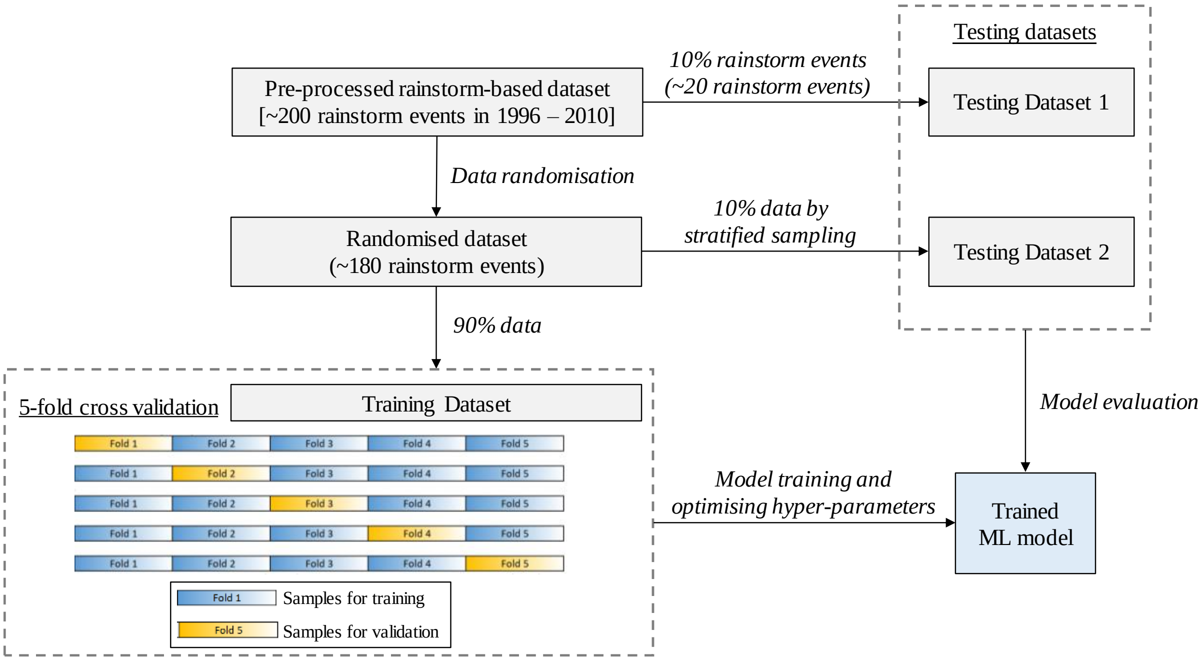



**Figure 8:** *Dataset resampling for model training and testing*

## 5 Conclusion

Riding on the comprehensive landslide-related data inventories collected and maintained over the years, the GEO has been exploring the potential of improving the current landslide prediction models with the use of artificial intelligence. Promising progress and outcome are obtained from the exploratory study of the two existing landslide prediction models so far.

The pilot study of the natural terrain landslide susceptibility model (Model A) suggests the use of ML technique coupling with big data analytics could bring about a 2-3 orders of magnitude enhancement to the resolution of the natural terrain landslide susceptibility model without compromising the prediction accuracy of the model. In this study, various aspects of ML including the workflow, characteristics of algorithms, feature selection as well as methods of model performance evaluation had been thoroughly investigated with respect to the modelling approach adopted. Emphasis was placed in each step on the importance of understanding the model and introducing sound domain engineering knowledge to ensure the derived ML model to remain physically meaningful. It is believed that the discussed approach and methodology can be extended to the development of Model A to a territory-wide scale for it to better facilitate the formulation of the slope safety management strategy and land use planning.

With reference to the above, the GEO has kick-started another exploratory study to apply ML to enhance the prediction model of landslides on man-made slopes. It is expected to gain insight into the relevance of various factors to the landslide occurrence, while the consistency with our domain knowledge on the landslide occurrence would be carefully assessed throughout the study. The upcoming technical development work would indubitably uncover the potential of ML in landslide prediction, to make a remarkable contribution to the slope safety system in Hong Kong.

## 6 Declarations

### 6.1 Acknowledgements

This paper is published with the permission of the Head of the Geotechnical Engineering Office and the Director of Civil Engineering and Development, the Government of the Hong Kong Special Administrative Region, China.

### 6.2 Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### How to Cite

Cheung *et al.* (2023). Exploratory Study of using Artificial Intelligence for Landslide Predictions. *AIJR Proceedings*, 215-229. https://doi.org/10.21467/proceedings.159.17

### References

Brabb, E. E. (1984). Innovative approaches to landslide hazard and risk mapping. *Proceedings of 4th International Symposium on Landslides*, *1*, 307–324. https://pubs.er.usgs.gov/publication/70197529

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and Regression Trees by Leo Breiman. In *Chapter 7* (Issue January). https://books.google.com/books/about/Classification_and_Regression_Trees.html?id=JwQx-WOmSyQC

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Cheung, R. W. M. (2021). Landslide risk management in Hong Kong. *Landslides 2021* , *18*(10), 3457–3473. https://doi.org/10.1007/S10346-020-01587-0

Chung, P. W., So, S. T., & Chu, E. K. (2022). Landslip Warning System in Hong Kong - Over 40 Years of Evolution. *Current Chinese Science*, *3*(2), 123–140. https://doi.org/10.2174/2210298102666220607152443

Cortes, C., Vapnik, V., & Saitta, L. (1995). Support-vector networks. *Machine Learning* , *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Cox, D. R. (1958). The Regression Analysis of Binary Sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, *20*(2), 215–232. https://doi.org/10.1111/J.2517-6161.1958.TB00292.X

Ho, K. K. S., & Cheung, C. (2021). Human fingerprints in landslides-the importance of strategic landslide risk management. *13th International Symposium on Landslides (ISL2020)*, 1–40. https://www.issmge.org/uploads/publications/105/106/ISL2020-169.pdf

Ho, K. K. S., & Lau, T. M. F. (2008). The systematic landslide investigation programme in Hong Kong. *Landslides and Engineered Slopes. From the Past to the Future, Two Volumes + CD-ROM*, 243–248. https://doi.org/10.1201/9780203885284-21

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *31st Conference on Neural Information Processing System (NIPS)*, 1–9. https://github.com/Microsoft/LightGBM.

Ko, F. W. Y. (2018). *Terrain-based Landslide Frequency Map for Natural Terrain in Hong Kong The Government of the Hong Kong Special Administrative Region*. https://www.cedd.gov.hk/filemanager/eng/content_963/er340links.pdf

Kong, V. W. W., Kwan, J. S. H., & Pun, W. K. (2020). Hong Kong's landslip warning system—40 years of progress. *Landslides*, *17*(6), 1453–1463. https://doi.org/10.1007/S10346-020-01379-6

Lam, C. C. L., Mak, S. H., & Wong, A. H. T. (1998). A new slope catalogue for Hong Kong. In K. K. S. Ho, J. N. Kay, & K. S. Li (Eds.), *In Slope engineering in Hong Kong, Proceedings of 17th Annual Seminar of Geotech. Division,* (pp. 235–242). A. A. Balkema. https://books.google.co.in/books/about/Slope_Engineering_in_Hong_Kong.html?id=vWKWwgEACAAJ&redir_esc=y

Li, H. W. M. ;, Lo, F. L. C. ;, Wong, T. K. C. ;, Cheung, R. W. M., Wai, H., Li, M., Leung, F., Lo, C., Kwok, T., Wong, C., Wai, R., & Cheung, M. (2022). Machine Learning-Powered Rainfall-Based Landslide Predictions in Hong Kong—An Exploratory Study. *Applied Sciences* , *12*(12), 6017. https://doi.org/10.3390/APP12126017

Lo, F. L. C., Law, R. P. H., & Ko, F. W. Y. (2022). *Territory-wide Rainfall-based Landslide Susceptibility Analysis The Government of the Hong Kong Special Administrative Region*. https://www.cedd.gov.hk/filemanager/eng/content_1017/er364.pdf

Ma, Z., Mei, G., & Piccialli, F. (2021). Machine learning for landslides prevention: a survey. *Neural Computing and Applications* , *33*(17), 10881–10907. https://doi.org/10.1007/S00521-020-05529-8

Mak, S. H., Ma, T. M., Chan, T. P., & Chan, D. C. (2001). Slope Information System-An Indispensable Tool for Hong Kong Slope Safety Management. *Proceedings of the Fourteenth Southeast Asian Geotechnical Conference*, 855–861. http://hkss.ced.gov.hk

Martinello, C., Cappadonia, C., & Rotigliano, E. (2023). Investigating the Effects of Cell Size in Statistical Landslide Susceptibility Modelling for Different Landslide Typologies: A Test in Central–Northern Sicily. *Applied Sciences* , *13*(2), 1145. https://doi.org/10.3390/APP13021145

Ng, C. W. W., Yang, B., Liu, Z. Q., Kwan, J. S. H., & Chen, L. (2021). Spatiotemporal modelling of rainfall-induced landslides using machine learning. *Landslides* , *18*(7), 2499–2514. https://doi.org/10.1007/S10346-021-01662-0

Pun, W.K., Wong, A.C.W. & Pang, P.L.R. 1999. Review of Landslip Warning criteria 1998/1999 (Special Project Report No. SPR 4/99), Geotechnical Engineering Office, Civil Engineering and Development Department, Hong Kong, 77 p.
https://www.cedd.gov.hk/eng/publications/geo/geo-reports/index.html

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/H0042519

Spackman, K. A. (1989). Signal Detection Theory: Valuable Tools for Evaluating Inductive Learning. *Proceedings of the Sixth International Workshop on Machine Learning*, 160–163. https://doi.org/10.1016/B978-1-55860-036-2.50047-3

Wang, H., Zhang, L., Luo, H., He, J., Cheung, R. W. M., Wang, H., Zhang, L., Luo, H., He, J., & Cheung, R. W. M. (2021). AI-powered landslide susceptibility assessment in Hong Kong. *Engineering Geology*, *288*, 106103. https://doi.org/10.1016/J.ENGGEO.2021.106103

Yu, Y. F. (2004). *Correlations Between Rainfall, Landslide Frequency and Slope Information For Registered Man-Made Slopes*. https://www.cedd.gov.hk/filemanager/eng/content_299/er144.pdf