

Data-driven landslide forecasting: Methods, data completeness, and real-time warning

Te Xiao^a, Li-Min Zhang^{a,b,*}

^a Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong Special Administrative Region

^b HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China



ARTICLE INFO

Keywords:

Rain-induced landslides
Landslide forecasting
Landslide risk
Data-driven methods
Machine learning

ABSTRACT

Various data-driven methods, including empirical, statistical, and machine learning methods, have been developed to promptly forecast rain-induced landslides. Their abilities differ considerably in spatio-temporal landslide prediction and in handling datasets of varying qualities. A challenging issue that significantly hinders the applications of data-driven methods is the data incompleteness in most landslide inventories, particularly the lack of accurate landslide time that is a vital link between each landslide and its triggering rainstorm. This study systematically compares the performances of three categories of data-driven methods for landslide prediction and proposes a novel machine learning model featured by probabilistic landslide modelling for spatio-temporal landslide prediction. The integrated machine learning model can be developed on a realistic landslide database, regardless of whether the landslide timing information is known or not. It not only promptly predicts the spatio-temporal evolution of landslides during a rainstorm but also reliably characterises the factual landslide risk, which provides a powerful real-time decision-making tool for landslide early warning and risk management. The model is validated against the landslide incidents in Hong Kong in the past 35 years both spatially and temporally, and outperforms other data-driven models in both prediction ability and accuracy.

1. Introduction

Rain-induced landslides are among the most catastrophic natural hazards in mountainous areas worldwide, such as Italy, Japan, the Himalayan belt, and Southeast China (e.g., Nadim et al., 2006; Kirschbaum et al., 2015; Jiang et al., 2021; Tonini et al., 2022). As a severe rainstorm can impact a large area and trigger thousands of landslides in a short time, it is essential to manage regional landslide risks and issue timely landslide warnings. The most critical question concerned by decision-makers is “where and when would landslides occur amid a rainstorm?” Answering this question requires the prediction of spatio-temporal landslide occurrence. Many physically-based and data-driven methods have been developed for this purpose. Physically-based methods (e.g., Crosta and Frattini, 2003; Baum et al., 2010; Chen and Zhang, 2014; Shen et al., 2018; Zhou et al., 2019; Luo et al., 2022; Medina et al., 2021) predict the spatio-temporal development of landslides incorporating slope failure mechanisms but are often unsuitable for real-time applications because of computational efficiency issues. By contrast, data-driven methods are efficient if regional landslide data (e.

g., Herrera et al., 2018; Froude and Petley, 2018; Lin and Wang, 2018; Cheung, 2021) and relevant geo-information data are available.

Several categories of data-driven methods have been developed for predicting landslide occurrence, including empirical, statistical, and machine learning methods.

- (1) Empirical methods, such as methods of frequency ratio, certainty factor, weights of evidence, and index of entropy (e.g., Van Westen et al., 2003; Yalcin et al., 2011; Devkota et al., 2013; Hong et al., 2017), focus on spatial landslide prediction to identify landslide-prone areas, often referred to as landslide susceptibility mapping. They compare feature differences between landslides and non-landslides one feature after another, also known as bivariate analysis, and combine all the differences into an empirical landslide index. Due to the mathematical simplicity, empirical methods can be easily implemented in spreadsheets and were widely adopted before the 2000s.
- (2) Statistical methods are often developed for temporal landslide prediction with emphasis on the number of landslides as a

* Corresponding author at: Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong Special Administrative Region.

E-mail addresses: xiaote@ust.hk (T. Xiao), cezhang@ust.hk (L.-M. Zhang).

function of rainfall features. Rainfall thresholds (e.g., Aleotti, 2004; Guzzetti et al., 2008; Segoni et al., 2018; Jiang et al., 2021) and rainfall-landslide correlations (e.g., Finlay et al., 1997; Ko and Lo, 2016; Kong et al., 2020; Xiao and Zhang, 2020) are two typical statistical methods widely deployed in landslide early warning worldwide. Statistically, the former identifies a rainfall threshold between landslides and non-landslides in the form of an intensity (I)-duration (D) relation of a storm (e.g., $I = \alpha D^{-\beta}$), whereas the latter establishes a statistical correlation (e.g., $\log(P) = \alpha + \beta R$) between landslide frequency (P) and critical rainfall characteristics (R).

- (3) Machine learning methods are powerful in handling numerous features and non-linear relations between features and landslides, thus widely applied to landslide susceptibility mapping since the 2000s as a multivariate update of empirical methods (e.g., Dai and Lee, 2002; Frattini et al., 2010; Van Den Eekhaut et al., 2012; Goetz et al., 2015; Huang et al., 2020; Can et al., 2021; Wang et al., 2021; Aslam et al., 2022; Kainthura and Sharma, 2022; Su et al., 2022). Various machine learning algorithms can be adopted, from logistic regression, decision trees, neural networks, etc., to emerging deep learning algorithms such as convolutional neural networks. Merghadi et al. (2020) made a comparative review of the 11 most commonly-used machine learning algorithms in landslide susceptibility studies. In contrast to empirical methods, most susceptibility results from machine learning methods can be transformed to a unified occurrence probability of landslides for fair comparison among different models.

Separated spatial and temporal predictions from conventional data-driven methods do not satisfy the practical needs of landslide risk management. An emerging research direction is to integrate the advantages of statistical methods in temporal prediction and machine learning methods in spatial mapping to enable dynamic landslide susceptibility mapping (e.g., Samia et al., 2020; Xing et al., 2021; Tehrani et al., 2022) or spatio-temporal landslide prediction (e.g., Pradhan et al., 2019; Kim et al., 2021; Xiao et al., 2022). A unified machine learning framework (Fig. 1) has been proposed by Xiao et al. (2022) to achieve spatio-temporal integration, by highlighting the role of dynamic rainfall data from the static slope and geographic data (unchanged during a rainstorm) and connecting each landslide to its triggering rainstorm. However, not every landslide database is qualified to develop such a prediction model, as most landslide inventories record spatial locations but lack accurate landslide time. The lack of landslide timing information particularly hinders the temporal prediction of landslide occurrence. This is the main reason why the majority of machine learning studies stay on landslide susceptibility mapping, despite the rapid development of advanced algorithms.

The primary objectives of this study are: (1) to systematically compare the performance of three data-driven methods (i.e., empirical, statistical, and machine learning methods) in landslide prediction; and (2) to propose a novel machine learning model for spatio-temporal landslide prediction, with a probabilistic strategy for handling incomplete landslide data and applications to real-time landslide early warning. The challenging data incompleteness issue will be reviewed first, followed by the development of a new machine learning model that accommodates incomplete data. The model will be validated against the

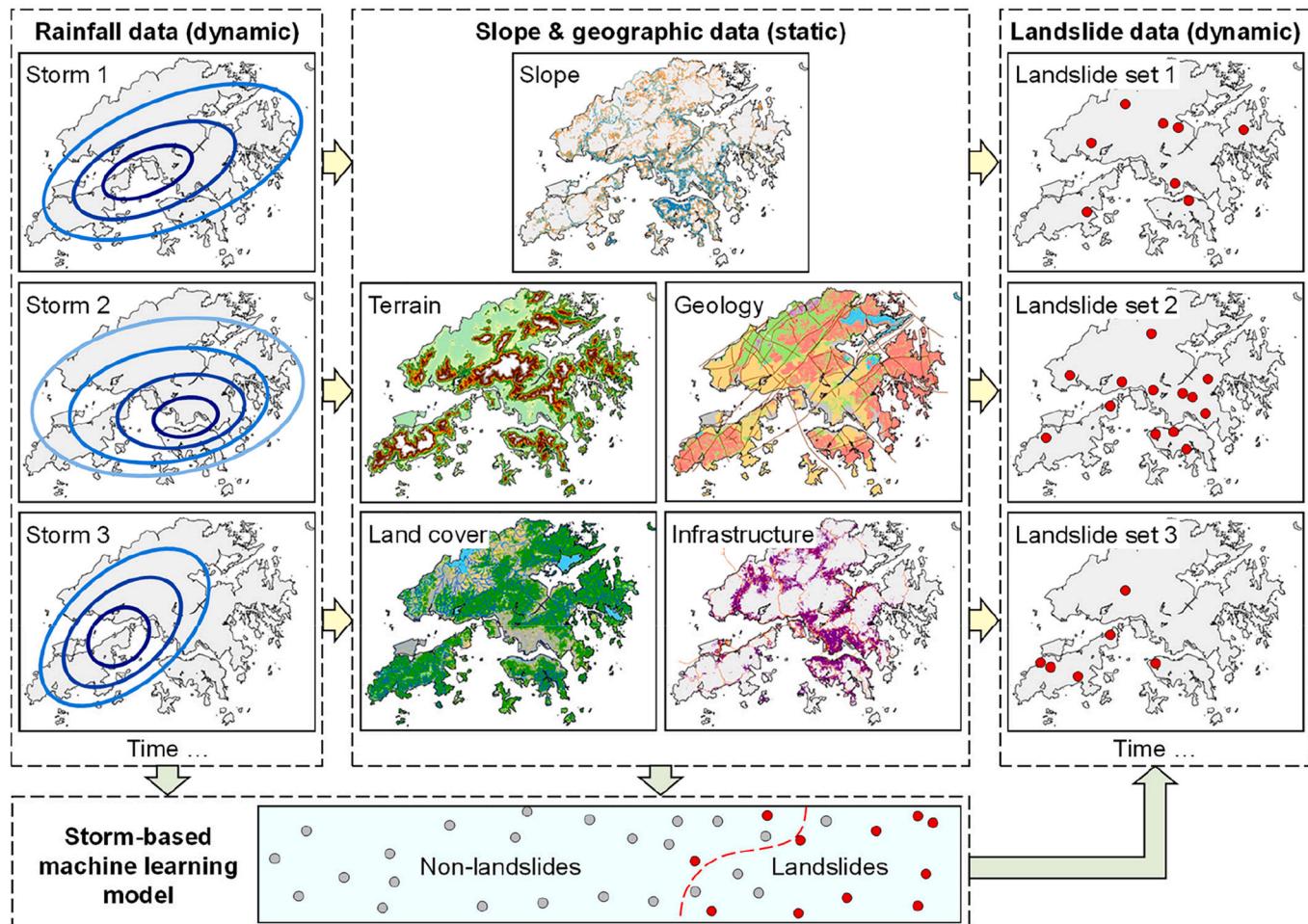


Fig. 1. Framework of spatio-temporal landslide forecasting (adapted from Xiao et al. (2022)).

landslide incidents in Hong Kong in the past 35 years both temporally and spatially and compared with other data-driven models. Finally, its application in real-time landslide early warning will be demonstrated.

2. Incomplete landslide data

2.1. Conventional data processing methods

Developing data-driven methods for rain-induced landslide prediction requires at least three sources of data: rainfall data, slope data (e.g., slope features for man-made slopes, geological and hydrological features for natural terrain hillslopes), and landslide data, as shown in Fig. 1. Data incompleteness has been a top issue hindering data-driven methods (e.g., Feng and Jimenez, 2015; Steger et al., 2017; Merghadi et al., 2020). It is possible to handle incomplete rainfall and slope data. For instance, rainfall features at locations of slopes are widely estimated from those at rain gauges using spatial interpolation techniques (e.g., Rahardjo et al., 2020); slope features with limited missing values can be fixed using imputation techniques (e.g., Little and Rubin, 2020). The challenge mainly comes from incomplete landslide data, particularly the missing of landslide timing information. Although landslide time is not an explicit feature in landslide forecasting, it is vital to link historical landslides with their triggering rainstorms. Unfortunately, most landslide inventories do not provide accurate landslide times, as they are mainly compiled based on geological surveys or interpretation of remote sensing images over a certain period (e.g., once in a year or even less frequent), from which only the range of possible landslide time can be determined. Fig. 2 shows an example of landslide interpretation in Hong Kong. Landslides are interpreted by comparing two aerial photos taken in a short period: one on 1 March 2008 with no landslide and the other on 25 July 2008 with numerous fresh landslides. Rainfall records show that eight storms occurred during this period, including a once-in-a-millennium storm on 6–9 June 2008. It is uncertain which specific rainstorm triggered which landslides; hence all landslides are labelled as ‘failure in 2008’ for simplicity. Public engagement in landslide reporting would be a practical means to collect accurate landslide timing information (e.g., Kocaman and Gokceoglu, 2019; Cheung, 2021).

Conventional data processing cannot handle such incomplete landslide data well. One straightforward choice that limits training data to a small portion regarding time-known data will underestimate the predicted number of landslides undoubtedly. On the contrary, a common method that assigns all time-unknown landslides to the largest storm in the period (e.g., Meusburger and Alewell, 2008; Gao et al., 2018; Li et al., 2022) will lead to over-prediction. In a simple scenario with known years of landslides, Ko and Lo (2016) proposed a practical year-

storm adjustment approach according to the rainfall frequency to approximately transform the year-based rainfall-landslide correlations into storm-based ones. However, this adjustment approach has not been rigorously verified and may not be extensible to more complex scenarios.

2.2. Landslide databases in Hong Kong

Past landslides in Hong Kong have been compiled into two databases by the Geotechnical Engineering Office (GEO) of the Hong Kong SAR Government. One is the Enhanced Natural Terrain Landslide Inventory (ENTLI) (e.g., Ko and Lo, 2016; Gao et al., 2021; Ju et al., 2022), including 90,000 relict natural terrain landslides and over 21,000 recent natural terrain landslides in 1924–2019. These landslides are interpreted from annual aerial photos so that the failure time is only accurate to a year. The other is the catalogue of slope failure incidents (e.g., Xiao et al., 2022; Yang et al., 2022), which records 7933 man-made slope failures in 1984–2017. Prompt reporting from the public makes the failure time of man-made slopes almost accessible, accurate to a day in most incidents and sufficient to identify their triggering storms. This study will focus on man-made slope failures to ensure that landslide forecasting performance is verifiable in each storm. The developed methods can be transferred to natural terrain landslides in the future.

In the context of Hong Kong, man-made slopes refer to cut slopes, fill slopes, and retaining walls. A total of 59,763 man-made slopes had been registered by the GEO as of 2019 (Fig. 3) and suffered 3588 failures (out of the 7933 incidents) in 1984–2017 (Fig. 4): 2127 were triggered by 419 major storms in this period, 830 did not have accurate failure times but years of failures are known, and 631 were induced by other triggers (not considered). The 2127 and 830 landslides are referred to as storm-labelled (i.e., complete) and year-labelled (i.e., incomplete) landslides, respectively. A year-labelled landslide is a special case of a period-labelled landslide. The year 2008 had the maximum number of landslides, with 320 storm-labelled and 87 year-labelled. The record-breaking 6–9 June 2008 storm alone triggered 162 confirmable man-made slope failures.

To develop a storm-based landslide forecasting model (Fig. 1), 43 features are extracted as listed in Table 1 to create a record of landslide or non-landslide: 29 static features regarding slope information, terrain, geological condition, land cover, location, and annual rainfall (i.e., SF1–AR), and 14 dynamic features regarding maximum rolling rainfall amounts and cumulative antecedent rainfall amounts (i.e., R1h-A15d). Landslide time is not a learning feature but is a bridge linking a landslide to the rainfall features of its triggering storm. For example, slope 6SW-C/CR797 only had one failure during the 419 storms, forming 418

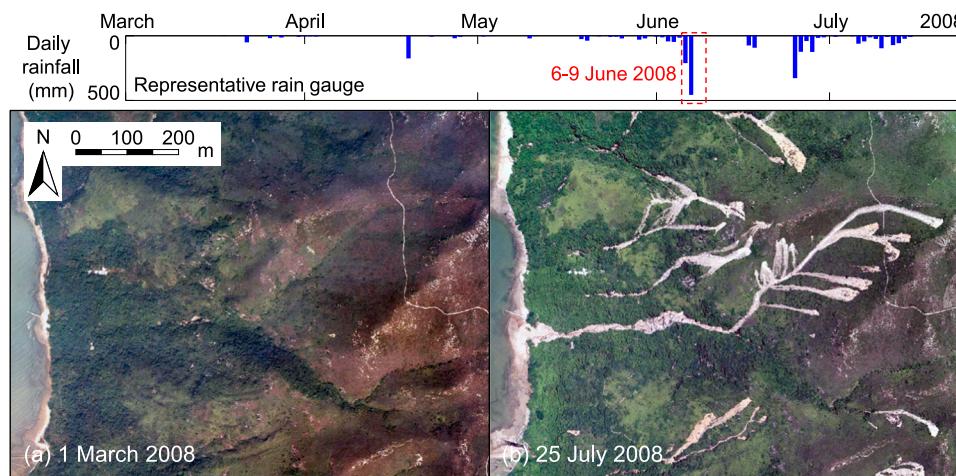


Fig. 2. Landslide interpretation in Hong Kong: (a) pre-landslide image on 1 March 2008; (b) post-landslide image on 25 July 2008 (source: www.hkmapservice.gov.hk).

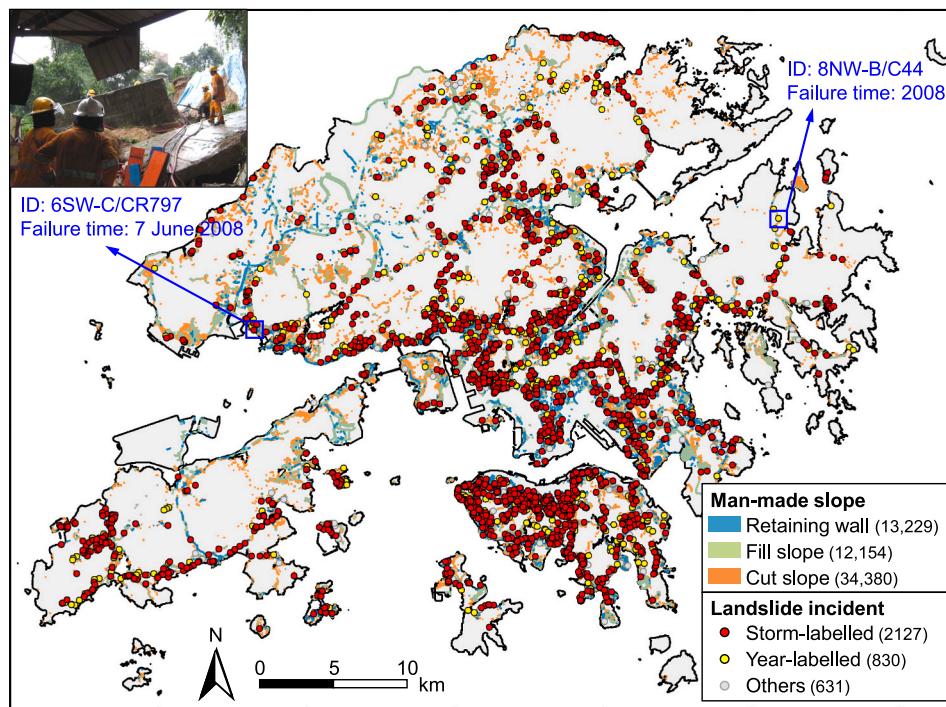


Fig. 3. Spatial distributions of man-made slopes and landslide incidents.

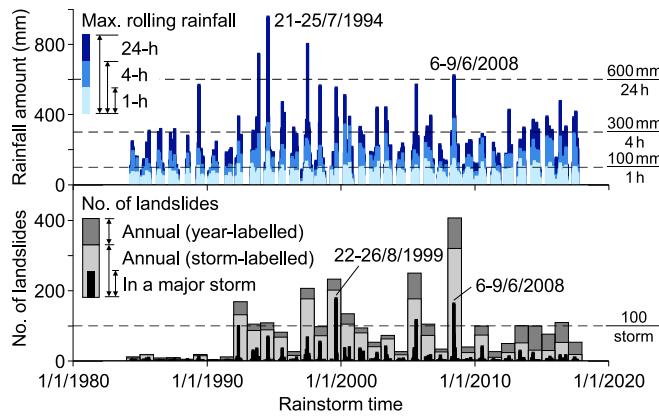


Fig. 4. Time series of major rainstorms and landslides in 1984–2017.

non-landslide records and one landslide record corresponding to the 6–June 2008 storm (Fig. 3). Details on the comprehensive databases and feature extraction can be referred to Xiao et al. (2022). The specific type of slope failures, e.g., sliding, washout, rockfall, and wall failure, is not considered in this study.

3. Machine learning-based landslide forecasting

3.1. Novel modelling strategy with incomplete landslide data

Among the three data-driven methods, machine learning methods are of the highest flexibility in handling various incomplete data. If all landslides are storm-labelled, referred to as a storm-labelled scenario, landslides and triggering storms can be connected through storm-based data integration (Fig. 5(a)). In such a complete data case, the N_l slopes experiencing N_s storms result in a total of $N = N_l \times N_s$ records. For each record, the failure probability p_s of a slope in the s th storm can be estimated by a machine learning algorithm, taking the logistic regression as an example, as:

$$p_s = \frac{1}{1 + \exp(-\mathbf{x}_s^T \boldsymbol{\theta})} \quad (1)$$

where \mathbf{x}_s is a set of 43 learning features (Table 1) corresponding to the s th storm and a constant term of unity; and $\boldsymbol{\theta}$ is a 44×1 vector of model parameters, which can be obtained by minimising a cost function $J(\boldsymbol{\theta})$ defined as:

$$J(\boldsymbol{\theta}) = -\frac{1}{N} \sum_{i=1}^N [y_i \ln p_i + (1 - y_i) \ln(1 - p_i)] \quad (2)$$

where y_i is a storm-based observation of the i th record: $y_i = 0$ for a non-landslide record and 1 for a landslide record; and p_i is the predicted failure probability of the i th record using Eq. (1).

Not all landslides are ideally storm-labelled. With respect to an incomplete period-labelled scenario where all landslides are labelled over a consistent period (e.g., one specific year for year-labelled landslides), it is conservative to claim that, for a particular landslide, the slope has at least one failure triggered by storms in this period. Therefore, the N_l slopes produce a total of $N = N_l \times N_p$ records after N_p -period observations (Fig. 5(a)). For each record, the probability P of observing at least one failure for a slope, after suffering n_s storms in a period, can be evaluated according to the probability theory as:

$$P = 1 - \prod_{s=1}^{n_s} (1 - p_s) \quad (3)$$

where p_s is the failure probability of the slope in the s th storm and can be calculated using a storm-based basic model like Eq. (1). Eq. (3) is referred to as the ‘at least one failure’ modelling strategy for incomplete landslides. Consequently, the cost function can be updated as:

$$J(\boldsymbol{\theta}) = -\frac{1}{N} \sum_{i=1}^N [y_i \ln P_i + (1 - y_i) \ln(1 - P_i)] \quad (4)$$

where y_i is a period-based observation of the i th record; and P_i is the predicted failure probability of the i th record using Eq. (3).

Note that Eq. (4) is similar to Eq. (2), but the landslide observation

Table 1

Features used for data-driven landslide forecasting.

Feature category	ID	Variable	Meaning
Man-made feature	1	SF1	Fill slope
	2	SF2	Cut slope
	–	[SF3]	Retaining wall
Formation time	3	FT	Post-1978
	–	[FT2]	Pre-1977
Slope material	4	SM1	Soil
	5	SM2	Rock
Slope cover	6	SC1	Percent of vegetation (%)
	7	SC2	Percent of chunam, shotcrete or other covers (%)
Slope characteristics	–	[SC3]	Percent of bare surface (%)
	8	SH	Slope height (m)
	9	SA	Slope angle (°)
	10	SB	Berms in slope
	11	SW	Weep holes in slope
Wall characteristics	12	SD	Number of drainages in slope
	13	WH	Wall height (m)
	14	WA	Wall angle (°)
	15	WB	Berms in wall
	16	WW	Weep holes in wall
Terrain	17	WD	Number of drainages in wall
	18	TE	Terrain elevation (m)
	19	TI	Terrain inclination (°)
Geological condition	20	GEO1	Granitic rock
	21	GEO2	Volcanic rock
	22	GEO3	Superficial deposit
	–	[GEO4]	Other geological conditions
Land cover	23	LC1	Cropland, shrub or grass
	24	LC2	Forest
Location	–	[LC3]	Other land covers
	25	DF	Distance to faults (0.1 km)
	26	DW	Distance to waterbodies (0.1 km)
	27	DR	Distance to roads (0.1 km)
	28	DB	Distance to buildings (0.1 km)
Annual rainfall	29	AR	Annual rainfall amount (mm*12 months)
	30	R1h	Maximum rolling 1-h rainfall amount (mm)
	31	R2h	Maximum rolling 2-h rainfall amount (mm)
	32	R4h	Maximum rolling 4-h rainfall amount (mm)
	33	R8h	Maximum rolling 8-h rainfall amount (mm)
	34	R12h	Maximum rolling 12-h rainfall amount (mm)
	35	R18h	Maximum rolling 18-h rainfall amount (mm)
	36	R24h	Maximum rolling 24-h rainfall amount (mm)
	37	R36h	Maximum rolling 36-h rainfall amount (mm)
	38	R48h	Maximum rolling 48-h rainfall amount (mm)
Antecedent rainfall	39	A2d	Cumulative 2-day antecedent rainfall amount (mm)
	40	A4d	Cumulative 4-day antecedent rainfall amount (mm)
	41	A7d	Cumulative 7-day antecedent rainfall amount (mm)
	42	A10d	Cumulative 10-day antecedent rainfall amount (mm)
	43	A15d	Cumulative 15-day antecedent rainfall amount (mm)

frequency has changed from every storm to a period consisting of several storms. Through the ‘at least one failure’ strategy, the storm-based failure probabilities can be transformed into period-based failure probabilities to match the period-based observations. In other words, although only $N_l \times N_p$ observations are available in a period-labelled scenario, Eq. (4) still computes $N_l \times N_s$ failure probabilities (where N_s is the total number of storms in all periods), consistent with that (i.e., Eq. (2)) in a storm-labelled scenario. This enables storm-based landslide forecasting using incomplete period-labelled data. Such a strategy is based on a solid probability theory and serves the same purpose as the year-storm adjustment approach proposed by Ko and Lo (2016). Comparing Eqs. (2) and (4), a non-landslide record in a period-labelled

scenario (i.e., $N = 1$ and $y = 0$ in Eq. (4)) is mathematically equivalent to N_s non-landslide records in a storm-labelled scenario (i.e., $N = N_s$ and $y_i = 0$, $i = 1, 2, \dots, N$, in Eq. (2)), which coincides with common sense. The ‘at least one failure’ strategy only updates the modelling of landslide records, relaxing from one certain storm to a period of several storms.

Furthermore, the ‘at least one failure’ strategy can be smoothly applied to any basic probability-based machine learning algorithms, as long as a storm-based landslide forecasting model can be developed following the same format of Eq. (1). Compared with other data-driven methods, machine learning methods have a compelling advantage: they can start with an initial storm-based model to realise forward landslide prediction and fit landslide observations iteratively by optimising model parameters, regardless of whether the observations are storm-labelled or period-labelled. To find the optimal model parameters, gradient-based methods such as the Newton-Raphson method (e.g., Nocedal and Wright, 2006) can be used to efficiently minimise the cost function as:

$$\theta^{(j+1)} = \theta^{(j)} - \alpha [\nabla^2 J(\theta^{(j)})]^{-1} \nabla J(\theta^{(j)}) \quad (5)$$

where $\theta^{(j)}$ is the parameter estimate at the j th iteration; α is the step size of backtracking line search satisfying the Armijo condition; and $\nabla J(\theta)$ and $\nabla^2 J(\theta)$ are the gradient vector and Hessian matrix of the cost function, respectively, and their closed-form solutions associated with the ‘at least one failure’ strategy are presented in Appendix A.

3.2. Integrated machine learning using realistic landslide data

Previous storm-labelled and period-labelled scenarios assume consistent completeness on failure times among all landslide data. The realistic landslide database is more often a complex mixture of various general cases, and the observation period varies from one slope to another. Fig. 5(b) gives a realistic example of 14 landslide records: nine failures are confirmed to be triggered by specific storms, and five are assigned to diverse periods. It is necessary to develop a landslide forecasting model using all available data, whether with or without accurate failure time, to maximise the value of information. Taking Fig. 5(b) as an example (i.e., $N_l = 8$ slopes, $N_s = 22$ storms, 9 storm-labelled landslides, and 5 period-labelled landslides), an integrated model can be established using the proposed ‘at least one failure’ strategy as follows:

- (1) Determine the possible triggering storms for each period-labelled landslide according to engineering judgement, such as storms between the last confirmation time of slope status and the reporting time of slope failure. Specifically, if a slope had failed in the 6–9 June 2008 storm (confirmed), another landslide incident reported on 1 July 2008 regarding the same slope means the slope might fail again between 10 June and 1 July, which covers two possible storms. Landslides without any time information can be treated as period-labelled landslides covering the entire extent.
- (2) Integrate all slopes one storm after another, leading to provisional $N_l \times N_s$ records (i.e., 176), and substitute the n_s uncertain records of each period-labelled landslide (i.e., $n_s = 2, 5, 7, 3$, and 22; 39 in total) with one failure record using the ‘at least one failure’ strategy, forming $N_{sl} = 176 - 39 = 137$ storm-labelled records and $N_{pl} = 5$ period-labelled records, as shown in Fig. 5(b).
- (3) Calculate the estimates of cost J , gradient ∇J , and Hessian matrix $\nabla^2 J$ for all storm-labelled records, E_{sl} , according to Eq. (2) and for all period-labelled records, E_{pl} , according to Eq. (4). The final estimate E can be derived in a weighted format as:

$$E = \frac{N_{sl}E_{sl} + N_{pl}E_{pl}}{N_{sl} + N_{pl}} \quad (6)$$

- (4) Optimise the model parameters according to the estimates from Eq. (6) to obtain an integrated storm-based model.

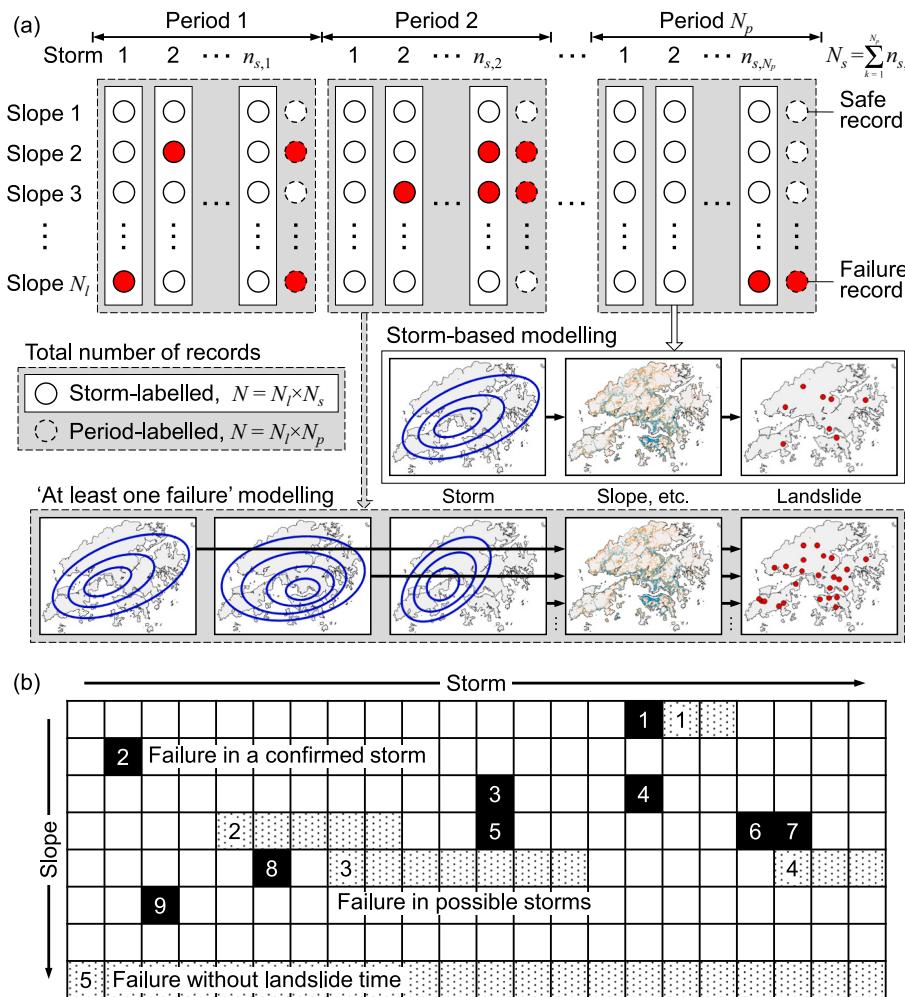


Fig. 5. Landslide data integration: (a) storm-labelled and period-labelled scenarios; (b) example of realistic scenario.

The above procedure treats all non-landslide records as storm-labelled records, due to the non-landslide equivalence between the two modelling strategies. The integrated model can be developed successfully regardless of whether the landslide time is known or not.

4. Benchmarking storm-based forecasting models

4.1. Comparison of data-driven methods

This section first benchmarks the three data-driven methods for landslide forecasting with complete storm-labelled data, namely the 2127 man-made slope failures in 1984–2017. The index of entropy (Devkota et al., 2013), rainfall-landslide correlations (Xiao and Zhang, 2020), and logistic regression (Xiao et al., 2022) are taken as representatives of empirical, statistical, and machine learning methods, respectively. For spatio-temporal landslide forecasting, the storm-based data integration (Xiao et al., 2022) is adopted to duplicate static features 419 times to match the dynamic rainfall features in all storms, leading to a total of 22 million records (i.e., valid records among 59,763 slopes \times 419 storms). Empirical and machine learning methods fully utilise all features listed in Table 1, while the statistical method only makes use of man-made features, slope materials, and maximum rolling 24-h rainfall amount. Besides, empirical and statistical methods need to discretise all features into multiple intervals, which may lose model generalisation ability to a certain degree. For example, if there are no historical rainfall data in an interval, such as 600–700 mm, they cannot forecast landslides when future rainfall falls in this interval. Computationally, training

efforts for empirical and statistical methods are almost negligible, and all data-driven methods are qualified for real-time prediction.

Fig. 6 compares the storm-based observations and predictions using the three data-driven methods. Statistical and machine learning methods can estimate the total number of landslides in each storm by summing the predicted failure probabilities of all slopes. By contrast, the empirical method only provides a symbolic landslide index, which is hardly applied to subsequent analyses like risk assessment. The three

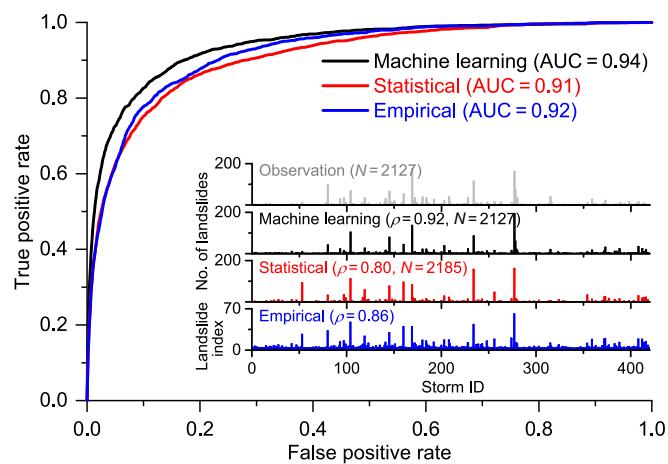


Fig. 6. Comparison of three data-driven methods.

methods have similar receiver operating characteristic curves, with areas under the curve (AUC) all higher than 0.9. The correlation coefficients between observations and predictions are 0.92, 0.80 and 0.86 for the machine learning, statistical and empirical methods, respectively. These results indicate that the machine learning method is the best among the three data-driven methods. It will be taken as a benchmark model to validate the storm-based landslide forecasting using the ‘at least one failure’ strategy when involving incomplete landslide data.

4.2. Validation of ‘at least one failure’ strategy

To validate the ‘at least one failure’ strategy, two incomplete scenarios are designed based on the 2127 storm-labelled landslides:

- (1) *Year-labelled scenario*. As a common case of period-labelled scenario, the accurate landslide time is deliberately ignored, but only the year of failure is used. Multiple failures of identical slopes in the same year are merged, leading to 2101 landslide records, and the total number of records is about 1.7 million (i.e., valid records among 59,763 slopes \times 34 years).

- (2) *No-failure-time scenario*. As an extreme case of period-labelled scenario, all time information is abandoned. The number of landslides reduces to 1869 in 34 years among a total of 59,428 observation records (i.e., 59,428 slopes formed in 1984–2017).

Models developed using the ‘at least one failure’ strategy are named ALOF-1Y and ALOF-34Y, respectively, for the two scenarios. Two

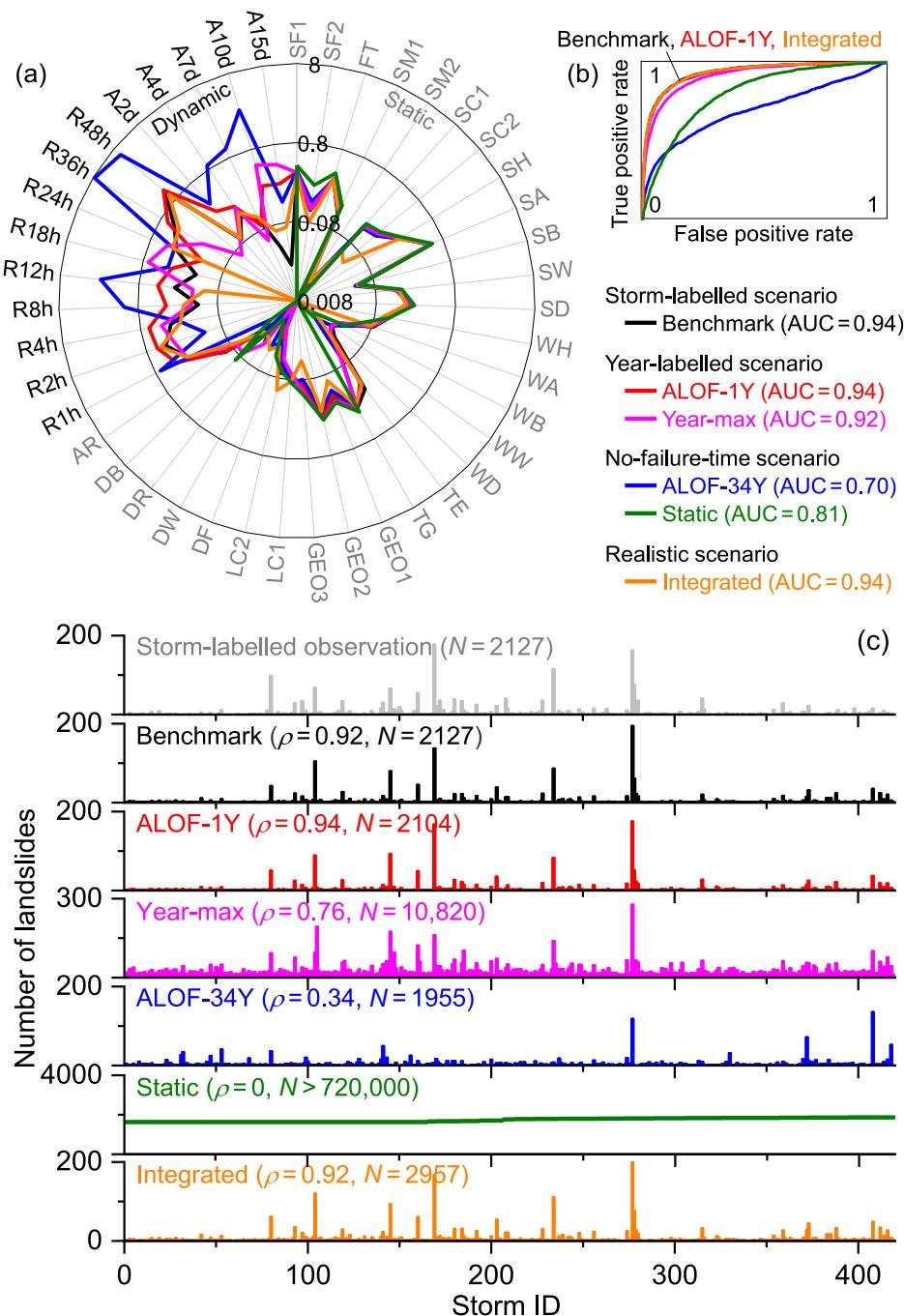


Fig. 7. Comparison of machine learning models: (a) normalised model parameters (notation in Table 1); (b) receiver operating characteristic curves; (c) number of landslides in storms.

simplified models are developed for comparison: a year-max model for the year-labelled scenario and a static model for the no-failure-time scenario. The year-max model assumes that all landslides in a year are triggered by the largest storm in each year in terms of the maximum rolling 24-h rainfall. The static model simply ignores dynamic rainfall features, thus does not need the failure time of landslides, and is equivalent to a conventional landslide susceptibility model. All models take the logistic regression as the basic prediction algorithm.

Fig. 7(a) plots a radar chart of the normalised model parameters in different models. The parameters corresponding to static features (i.e., SF1-AR) are almost consistent. This is expected as the static features are independent of storms and landslide times. Regarding the parameters of dynamic features (i.e., R1h-A15d), the ALOF-1Y model agrees well with the benchmark model, but the ALOF-34Y and year-max models deviate considerably. As shown in **Fig. 7(b)**, the AUC values for the benchmark, ALOF-1Y, and year-max models are close and high, but the ALOF-34Y model is even worse than the static model. Their landslide predictions in all 419 storms are compared in **Fig. 7(c)**. The benchmark and ALOF-1Y models provide similar landslide sequences compared to the observation, with high correlation coefficients between the predicted and observed numbers of landslides as 0.92 and 0.94, respectively. This validates the feasibility of ‘at least one failure’ strategy in a year-labelled scenario. On the contrary, the ALOF-34Y model mis-predicts landslide peaks with a small correlation coefficient of 0.34. The year-max model has a moderate correlation coefficient of 0.76, but significantly overestimates the number of landslides, particularly in small storms. This is not surprising since it concentrates all landslides induced by small storms into one largest storm. The bias expands notably if one year includes several storms with similar intensities. The static landslide susceptibility model without dynamic rainfall features predicts a steady number of landslides, which is not suitable for temporal landslide forecasting.

The most remarkable finding is that the ALOF-1Y model based on ‘at least one failure’ modelling, knowing only the year of failure, has the same high accuracy in storm-based landslide forecasting as the benchmark model knowing accurate failure time. This significantly relaxes the requirement on landslide timing and enhances the application of storm-based landslide forecasting with many time-unknown landslide inventories worldwide, such as the ENTLI in Hong Kong. The effectiveness of ‘at least one failure’ strategy relies on the observation frequency of landslides, which will be further elaborated in the following section.

4.3. Impact of landslide observation frequency

The frequencies of landslide observation are every storm, once in a year, and once over 30 years in the storm-labelled, year-labelled, and no-failure-time scenarios, respectively. To explore the impact of landslide observation frequency, four additional models are further developed with landslide observation intervals of 2, 5, 8, and 10 years and the same ‘at least one failure’ strategy. **Fig. 8** presents the variation of model performance with a decreasing frequency of landslide observation. The model performance drops dramatically when the observation frequency is lower than once in eight years. Fewer landslide peaks can be correctly identified as the time interval increases. According to the predicted landslide sequence, the observation frequency should be higher than once in five years, or better be annually or biennially. This advocates more frequent landslide surveys to minimise the uncertainty in landslide timing. An appropriate frequency should be region-specific and depend on how many major storms are expected in one year. The sensitivity analysis indicates that annual or biennial landslide survey is sufficiently frequent for regions like Hong Kong.

4.4. Identification of triggering storms

Assuming that all landslides in a period are triggered by the maximum rainstorm in this period (i.e., year-max model) is proven to

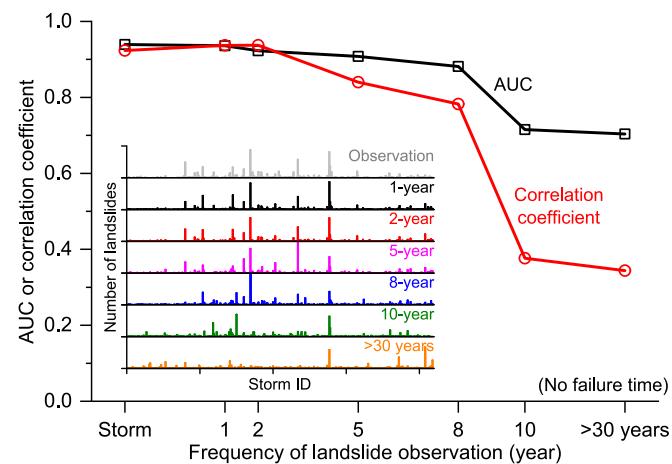


Fig. 8. Impact of landslide observation frequency.

over-predict landslides. How to determine the triggering storm of each time-unknown landslide remains an open question. The well-worked ALOF-1Y model provides an ideal tool to determine the triggering storm in a year-labelled scenario. Take the 320 storm-labelled landslides in 2008 as an example. They were distributed in 11 of 16 major storms in 2008, as illustrated in **Fig. 9**. The four most severe storms, i.e., 6–9 June, 11–19 June, 25 June–1 July, and 6–17 July storms, induced 50.6%, 23.4%, 7.2%, and 10.9% of all observed landslides, respectively.

The ALOF-1Y model estimates the failure probabilities of each slope in all 16 storms. The normalised failure probability over all storms (i.e., $p_s / \sum_{s=1}^{n_s} p_s$ and $n_s = 16$) can be taken as the predicted landslide-triggering probability of the corresponding storm in return. Recall the

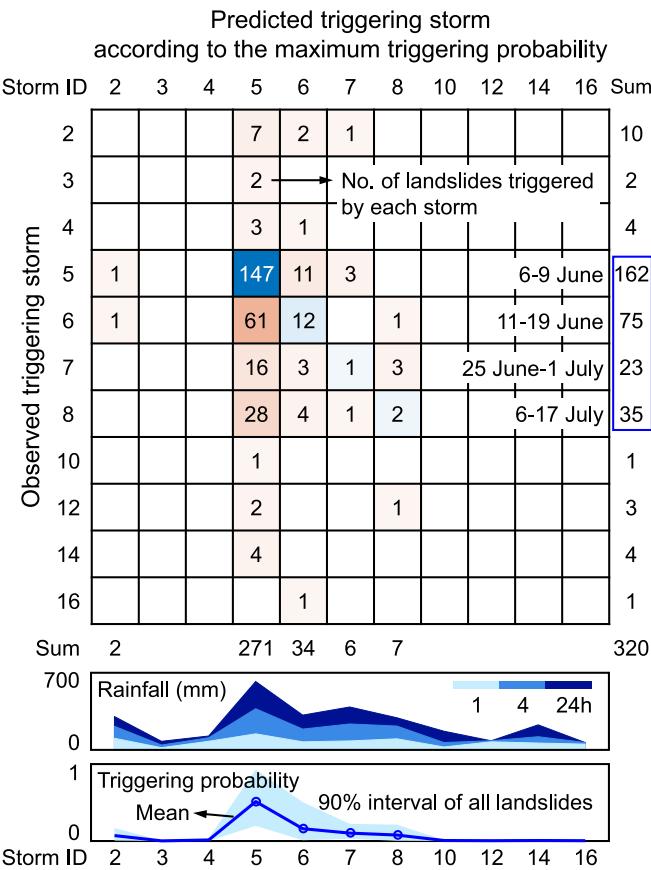


Fig. 9. Identification of triggering storms for time-unknown landslides in 2008.

slope example 6SW-C/CR797 in Fig. 3. Its predicted failure probability varies from 2.04×10^{-5} in the 13th storm to 1.80×10^{-3} in the 5th storm (i.e., the 6–9 June storm) and the total failure probability in all 16 storms is 4.80×10^{-3} . The 6–9 June storm has the highest triggering probability of 38%, and the following three storms correspond to 20%, 17%, and 10%, respectively. Repeating this process for all 320 failed slopes draws the distribution of landslide-triggering probabilities for all storms, as demonstrated in the bottom part of Fig. 9. The mean triggering probabilities of the four most severe storms are 51.4%, 16.1%, 10.3%, and 7.8%, respectively, which coincide with the observed proportions of landslides. If the storm with the highest triggering probability is taken as a triggering storm, as shown in the confusion matrix in Fig. 9, 85% of landslides will be crudely assigned to the 6–9 June storm, close to the allocation ratio (i.e., 100%) in a year-max model. In contrast to simply identifying a certain triggering storm according to the maximum rainfall amount or triggering probability, it is more rational to probabilistically allocate a landslide to every possible storm. The ‘at least one failure’ strategy does properly consider all possible storms with varying triggering probabilities, which makes the ALOF-1Y model outperforms the year-max model in a year-labelled scenario.

5. Real-time landslide forecasting for Hong Kong

5.1. Development of an integrated model

Considering the success of the ‘at least one failure’ strategy in a year-labelled scenario, the 830 year-labelled landslides (Fig. 3) will be utilised as well, together with the 2127 storm-labelled landslides, to develop an integrated storm-based landslide forecasting model for Hong Kong. Possible triggering storms for each time-unknown landslide are determined according to the failure reporting time recorded in the database to shrink the period from one year to a few candidate storms. The model parameters and predicted numbers of landslides of the integrated model are also plotted in Fig. 7. As the amount of storm-labelled data is 1.6 times larger than the year-labelled data, the benchmark model is sufficiently close to the integrated model in terms of all metrics. Fig. 10 summarises the landslide sequence year by year to further verify year-labelled landslides. The prediction of the integrated model highly coincides with the yearly total number of storm-labelled and year-labelled landslides (i.e., $\rho = 0.94$), and the difference between the integrated and benchmark models also correlates with the observed number of year-labelled landslides (i.e., $\rho = 0.82$). Both indicate that those time-unknown landslides can be properly modelled using the ‘at

least one failure’ strategy. Note that the proportion of year-labelled landslides has increased since 2000; particularly after 2011, the year-labelled landslides even account for more than half. This highlights the importance of considering time-unknown landslides to avoid underestimation. The integrated model opens such a possibility and can reliably characterise the factual landslide risk.

Fig. 11 visualises the rainfall-landslide relations in the integrated machine learning model and conventional statistical correlations (Xiao and Zhang, 2020). The results of the integrated model are mapped to the dimension of maximum rolling 24-h rainfall and averaged with different types of man-made slopes, namely cut slopes, fill slopes and retaining walls. The two models illustrate similar increasing trends in failure probability as the rainfall amount increases, with cut slopes having the highest failure probability, followed by fill slopes and retaining walls. The prescribed bilinear relations manually change gradients at the 24-h rainfall amount of 300 mm in the statistical model, which can be automatically captured by the integrated model. Moreover, the integrated model predicts considerably higher failure probabilities of cut slopes and retaining walls at extremely high rain intensities (i.e., > 450 mm / 24 h). To adapt to the increasing landslide risk, it is necessary to upgrade the landslide early warning system in Hong Kong from the statistical model to the integrated machine learning model.

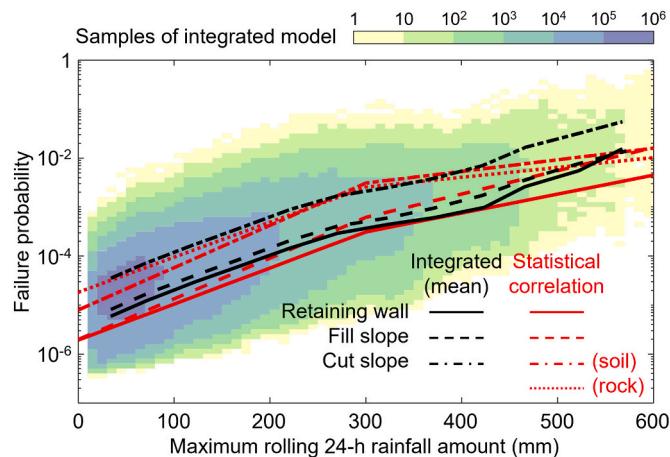


Fig. 11. Rainfall-landslide relations in machine learning and statistical models.

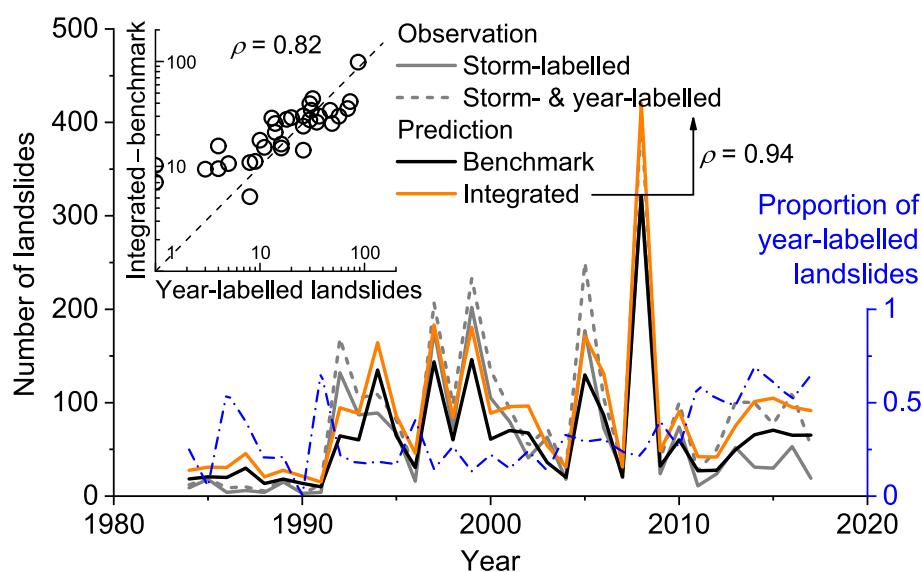


Fig. 10. Comparison of yearly total number of landslides.

5.2. Case study of landslides in 2008

The integrated landslide forecasting model is further applied to a case study of landslides in 2008 to demonstrate its capability for spatio-temporal landslide forecasting. The failure probabilities of all 59,763 slopes can be predicted instantly (i.e., < 1 s) after the real-time update of rainfall distribution. Fig. 12 presents the spatio-temporal evolution of rainstorms and predicted landslides during the 6–9 June 2008 storm. The rainfall intensity amplifies dramatically from 30 h to 34 h, with a maximum 4-h rainfall amount of 373 mm. Hardly any landslide occurs before 30 h, but numerous landslides burst shortly on western Lantau Island at 32 h and on north-western Hong Kong Island and western Kowloon at 34 h. Nearly 70% of landslides are triggered within the critical four hours. A landslide warning was issued by the GEO at 25 h of the storm, successfully reducing landslide losses. The predicted landslide distribution at 96 h with a total of 253 landslides agrees well with the observation (Fig. 13). The predicted number is greater than the confirmed number of landslides in this storm (i.e., 162) because 87 time-unknown landslides in 2008 are not assigned to any storm yet.

To accurately assess the effect of time-unknown landslides, all 16 major storms in 2008 should be analysed collectively. For the convenience of illustration, only the top 4 severe storms in 2008 are illustrated. According to the estimation of triggering probabilities in Section 4.4, these four storms are likely to have triggered the 87 time-unknown landslides, with a high mean probability of 88%. As shown in Fig. 13, the 11–19 June, 25 June–1 July, and 6–17 July storms additionally provide 76, 26, and 19 predicted landslides, respectively. Most observed landslides are located within the predicted landslide-prone areas. Even the

locations of individual time-unknown landslides can be well predicted using the integrated model, as marked in the bottom part of Fig. 13. The total predicted number of landslides (i.e., 374) in the four storms is very close to the sum of total confirmed landslides (i.e., 295) and time-unknown landslides (i.e., 87). The integrated model can well handle time-unknown landslides both spatially and temporally and provides a powerful real-time decision-making tool for landslide early warning and risk management.

6. Summary and conclusions

This study reviews and compares three data-driven methods for real-time landslide prediction, including empirical, statistical, and machine learning methods. To overcome the data completeness problem in most landslide inventories, a novel machine learning model is developed to make full use of time-unknown landslide data that are often abandoned in conventional methods. The model is validated against the landslide incidents in Hong Kong in the past 35 years both spatially and temporally. Major conclusions are as follows:

- (1) The machine learning method performs the best among the three data-driven methods, not only attaining the highest scores of all metrics, but also enabling the effective use of both complete and incomplete landslide data.
- (2) The ‘at least one failure’ strategy transforms the storm-based failure probabilities into period-based failure probabilities to match the observation of landslides in a certain period. It opens the possibility to use incomplete landslide data in storm-based

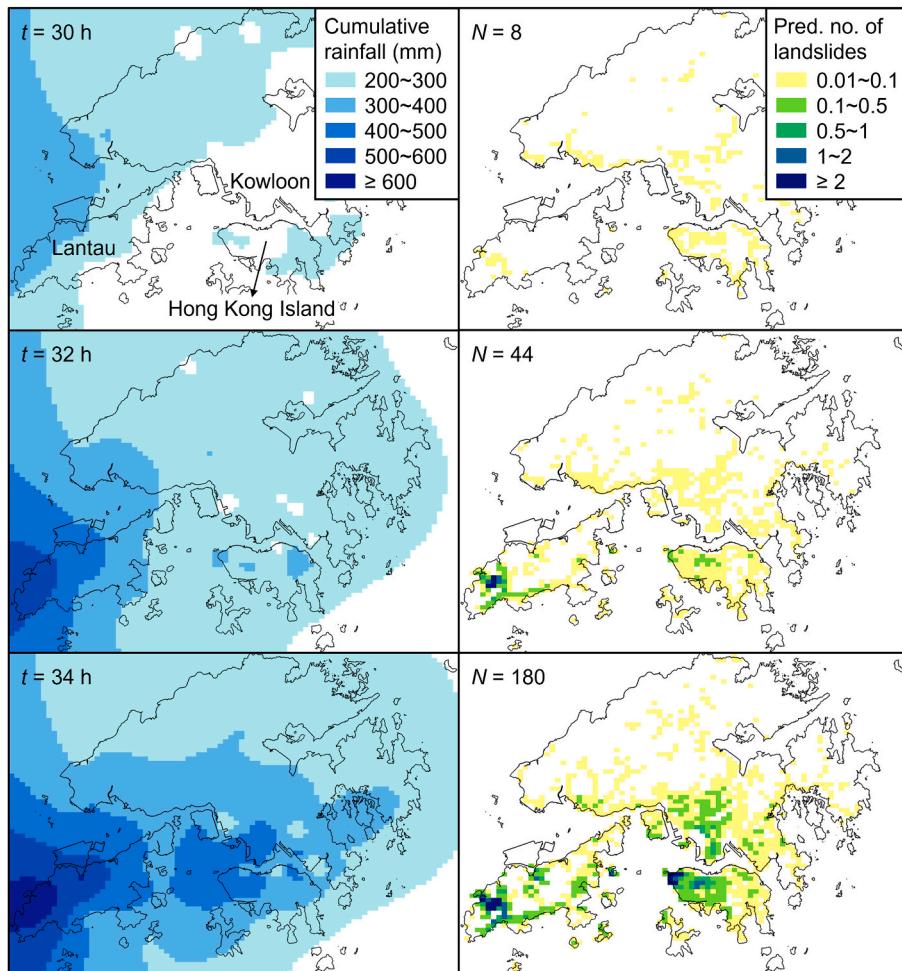


Fig. 12. Predicted spatio-temporal evolution of landslides in the 6–9 June 2008 storm.

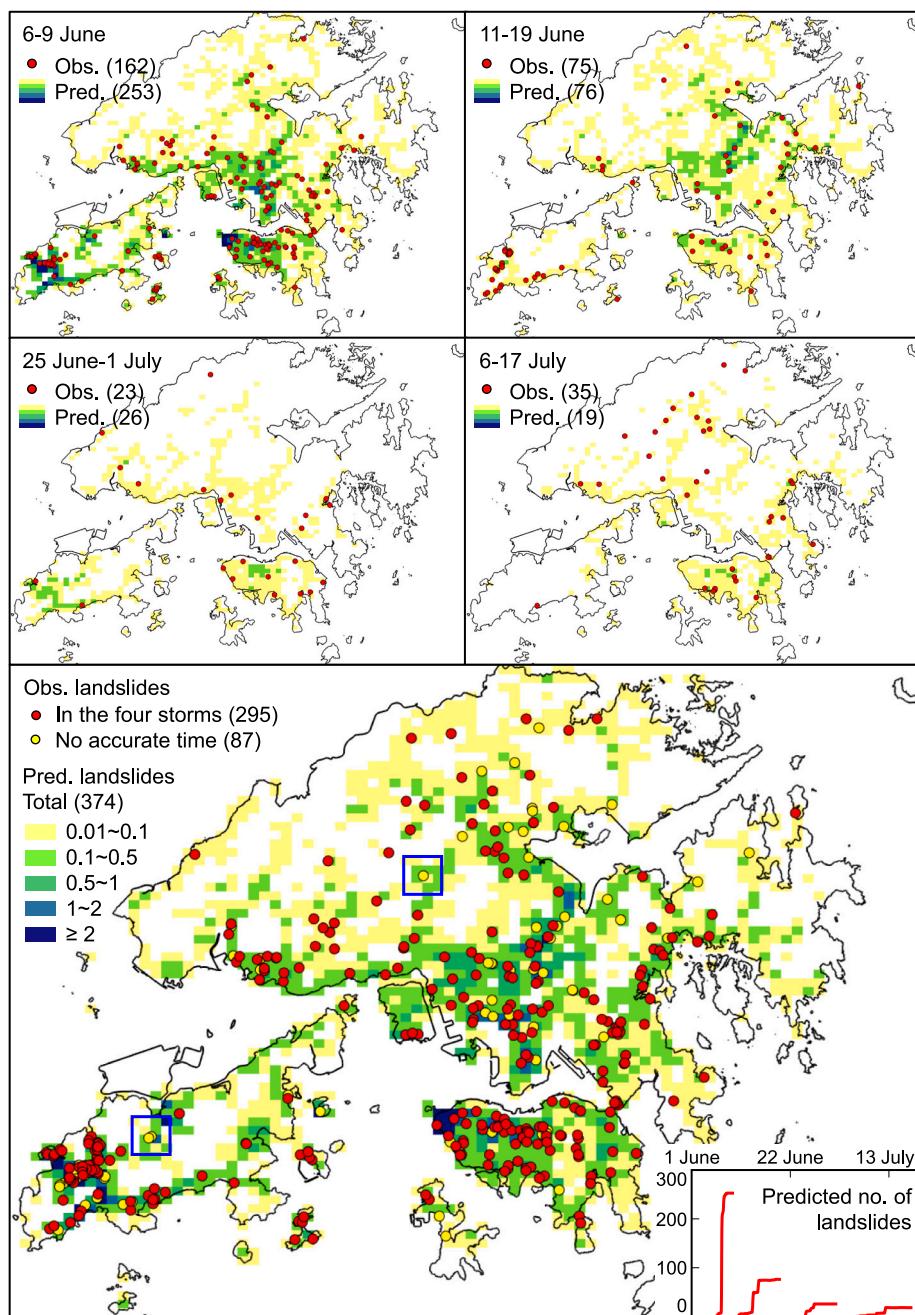


Fig. 13. Predicted landslides in the four most severe storms in 2008.

landslide forecasting and is applicable to any probability-based machine learning algorithms.

- (3) Applying the ‘at least one failure’ strategy, the ALOF-1Y model for landslides with information of only the year of failure has the same high prediction accuracy as the benchmark model for landslides with accurate failure time information. Annual or biennial landslide surveys will be sufficiently frequent for regions like Hong Kong.
- (4) The integrated machine learning model utilising both time-known and time-unknown landslides can reliably characterise the factual landslide risk. Compared with the statistical correlations used in a landslide early warning system, the integrated model predicts considerably higher failure probabilities of cut slopes and retaining walls at extremely high rain intensities.
- (5) The integrated model can promptly predict the spatio-temporal evolution of landslides as a rainstorm progresses, which

provides a powerful real-time decision-making tool for landslide early warning and risk management. The case study confirms that most observed landslides are located within the predicted landslide-prone areas and even the locations of individual time-unknown landslides can be well predicted.

CRediT authorship contribution statement

Te Xiao: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft. **Li-Min Zhang:** Funding acquisition, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the National Key R&D Program of China

(No. 2021YFC3001000), the Research Grants Council of the Hong Kong SAR (Nos. 16205719, 16203720, N_HKUST620/20, and AoE/E-603/18), and the Hetao Shenzhen-Hong Kong Science and Technology Innovation Cooperation Zone (No. HZQB-KCZYB-2020083). The authors acknowledge the data support from the Geotechnical Engineering Office of the Civil Engineering and Development Department, the Government of the Hong Kong SAR. All data used are open for public use by the Geotechnical Engineering Office (<https://www.geomap.cedd.gov.hk/GEOOpenData/eng/Default.aspx>).

Appendix A. Closed-form derivatives of Eq. (4)

If the logistic regression (Eq. (1)) is taken as the basic algorithm for storm-based landslide forecasting, the partial derivative of the probability of at least one failure in an n_s -storm period (Eq. (3)) with respect to the m th model parameter θ_m can be derived as:

$$\frac{\partial P}{\partial \theta_m} = \sum_{s=1}^{n_s} \left[\frac{\partial p_s}{\partial \theta_m} \prod_{k=1, k \neq s}^{n_s} (1 - p_k) \right] = \left[\prod_{s=1}^{n_s} (1 - p_s) \right] \left[\sum_{s=1}^{n_s} \frac{1}{1 - p_s} \frac{\partial p_s}{\partial \theta_m} \right] = (1 - P) \sum_{s=1}^{n_s} x_{sm} p_s \quad (\text{A.1})$$

where x_{sm} is the m th feature (corresponding to θ_m) in \mathbf{x}_s ; and $\partial p_s / \partial \theta_m = x_{sm} p_s (1 - p_s)$. Based on Eq. (A.1), the closed-form first-order and second-order partial derivatives of the updated cost function (Eq. (4)) can be derived, respectively, as:

$$\frac{\partial J}{\partial \theta_m} = \frac{1}{N} \sum_{i=1}^N \frac{P_i - y_i}{P_i(1 - P_i)} \frac{\partial P_i}{\partial \theta_m} = \frac{1}{N} \sum_{i=1}^N \left[\left(1 - \frac{y_i}{P_i} \right) \sum_{s=1}^{n_s} x_{ism} p_{is} \right] \quad (\text{A.2})$$

$$\frac{\partial^2 J}{\partial \theta_m \partial \theta_n} = \frac{1}{N} \sum_{i=1}^N \left[\frac{y_i(1 - P_i)}{P_i^2} \sum_{s=1}^{n_s} x_{ism} p_{is} \sum_{s=1}^{n_s} x_{isn} p_{is} + \left(1 - \frac{y_i}{P_i} \right) \sum_{s=1}^{n_s} x_{ism} x_{isn} p_{is} (1 - p_{is}) \right] \quad (\text{A.3})$$

where x_{ism} is the m th feature in the s th storm of the i th record. If one period contains only one storm (i.e., $n_s = 1$), $P_i = p_{is} = p_i$ and $x_{ism} = x_{im}$, hence Eqs. (A.2) and (A.3) can be simplified as:

$$\frac{\partial J}{\partial \theta_m} = \frac{1}{N} \sum_{i=1}^N x_{im} (p_i - y_i) \quad (\text{A.4})$$

$$\frac{\partial^2 J}{\partial \theta_m \partial \theta_n} = \frac{1}{N} \sum_{i=1}^N x_{im} x_{in} p_i (1 - p_i) \quad (\text{A.5})$$

These two equations are consistent with those in the conventional single storm-based logistic regression (i.e., derivatives of Eq. (2)).

With these closed-form solutions, it is easy to compute the gradient vector and Hessian matrix and use them in the cost minimisation using Eq. (5). For efficient implementation, loops for various storms, slopes, and model parameters can be vectorised. Consider an example that N slopes undergo n_s storms in one period, which leads to N observations using the ‘at least one failure’ strategy. Fig. A.1 demonstrates the vectorised implementation of the cost function (Eq. (4)), gradient (Eq. (A.2)), and Hessian matrix (Eq. (A.3)) in MATLAB. Similarly, the estimates of multiple periods can be summed in a weighted manner according to the number of observations in different periods.

```

function [J, grad, hess] = costFunction(theta, X, y)
% theta      : np*1 model parameters
% X          : N*np*ns features, including a column of 1
% y          : N*1 observations (0 = stable, 1 = failed)
% N, np, ns: Numbers of samples, model parameters,
%             and storms in one period, respectively

% Prediction
[N, np, ns] = size(X);
X3 = permute(X, [1, 3, 2]); % N*ns*np
X2 = reshape(X3, N*ns, np); % Nxns*np
Xt = reshape(X2*theta, N, ns); % N*ns
p = constrain( 1./(1+exp(-Xt)) ); % N*ns
P = constrain( 1-prod(1-p, 2) ); % N*1

% Cost
J = -sum(y.*log(P) + (1-y).*log(1-P))/N; % 1*1

% Gradient
Xp = permute(sum(X3.*p, 2), [1, 3, 2]); % N*np
grad = Xp'*(1-y./P)/N; % np*1

% Hessian matrix
w1 = y.* (1-P)./P.^2; % N*1
w2 = repmat(1-y./P, ns, 1); % Nxns*1
Xpw = Xp.*sqrt(w1); % N*np
Xw = X2.*sqrt(p(:).* (1-p(:))); % Nxns*np
hess = (Xpw'*Xpw + Xw'* (w2.*Xw))/N; % np*np

function p = constrain(p)
% Constrain between upper and lower limits
lower = eps; upper = 1-eps;
p(p<lower) = lower;
p(p>upper) = upper;

```

Fig. A.1. Vectorised implementation of the cost function in MATLAB.

References

- Aleotti, P., 2004. A warning system for rainfall-induced shallow failures. Eng. Geol. 73 (3–4), 247–265. <https://doi.org/10.1016/j.enggeo.2004.01.007>.
- Aslam, B., Zafar, A., Khalil, U., 2022. Comparative analysis of multiple conventional neural networks for landslide susceptibility mapping. Nat. Hazards 1–35. <https://doi.org/10.1007/s11069-022-05570-x>.
- Baum, R.L., Godt, J.W., Savage, W.Z., 2010. Estimating the timing and location of shallow rainfall-induced landslides using a model for transient, unsaturated infiltration. J. Geophys. Res. Earth Surf. 115 (F3), F03013. <https://doi.org/10.1029/2009JF001321>.
- Can, R., Kocaman, S., Gokceoglu, C., 2021. A comprehensive assessment of XGBoost algorithm for landslide susceptibility mapping in the upper basin of Ataturk dam, Turkey. Appl. Sci. 11 (11), 4993. <https://doi.org/10.3390/app11114993>.
- Chen, H.X., Zhang, L.M., 2014. A physically-based distributed cell model for predicting regional rainfall-induced shallow slope failures. Eng. Geol. 176, 79–92. <https://doi.org/10.1016/j.enggeo.2014.04.011>.
- Cheung, R.W.M., 2021. Landslide risk management in Hong Kong. Landslides 18 (10), 3437–3473. <https://doi.org/10.1007/s10346-020-01587-0>.
- Crosta, G.B., Frattini, P., 2003. Distributed modelling of shallow landslides triggered by intense rainfall. Nat. Hazards Earth Syst. Sci. 3 (1/2), 81–93. <https://doi.org/10.5194/nhess-3-81-2003>.
- Dai, F.C., Lee, C.F., 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42 (3–4), 213–228. [https://doi.org/10.1016/S0169-555X\(01\)00087-3](https://doi.org/10.1016/S0169-555X(01)00087-3).
- Devkota, K.C., Regmi, A.D., Pourghasemi, H.R., Yoshida, K., Pradhan, B., Ryu, I.C., Dhital, M.R., Althuwainee, O.F., 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. Nat. Hazards 65 (1), 135–165. <https://doi.org/10.1007/s11069-012-0347-6>.
- Feng, X., Jimenez, R., 2015. Predicting tunnel squeezing with incomplete data using Bayesian networks. Eng. Geol. 195, 214–224. <https://doi.org/10.1016/j.enggeo.2015.06.017>.
- Finlay, P.J., Fell, R., Maguire, P.K., 1997. The relationship between the probability of landslide occurrence and rainfall. Can. Geotech. J. 34 (6), 811–824. <https://doi.org/10.1139/t97-047>.
- Frattini, P., Crosta, G., Carrara, A., 2010. Techniques for evaluating the performance of landslide susceptibility models. Eng. Geol. 111 (1–4), 62–72. <https://doi.org/10.1016/j.enggeo.2009.12.004>.
- Froude, M.J., Petley, D.N., 2018. Global fatal landslide occurrence from 2004 to 2016. Nat. Hazards Earth Syst. Sci. 18 (8), 2161–2181. <https://doi.org/10.5194/nhess-18-2161-2018>.
- Gao, L., Zhang, L.M., Cheung, R.W.M., 2018. Relationships between natural terrain landslide magnitudes and triggering rainfall based on a large landslide inventory in Hong Kong. Landslides 15 (4), 727–740. <https://doi.org/10.1007/s10346-017-0904-x>.
- Gao, L., Zhang, L.M., Chen, H.X., Fei, K., Hong, Y., 2021. Topography and geology effects on travel distances of natural terrain landslides: evidence from a large multi-temporal landslide inventory in Hong Kong. Eng. Geol. 292, 106266 <https://doi.org/10.1016/j.enggeo.2021.106266>.
- Goetz, J.N., Brenning, A., Petschko, H., Leopold, P., 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput. Geosci. 81, 1–11. <https://doi.org/10.1016/j.cageo.2015.04.007>.
- Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P., 2008. The rainfall intensity–duration control of shallow landslides and debris flows: an update. Landslides 5 (1), 3–17. <https://doi.org/10.1007/s10346-007-0112-1>.
- Herrera, G., Mateos, R.M., García-Davalillo, J.C., et al., 2018. Landslide databases in the Geological surveys of Europe. Landslides 15 (2), 359–379. <https://doi.org/10.1007/s10346-017-0902-z>.
- Hong, H., Chen, W., Xu, C., Youssef, A.M., Pradhan, B., Tien Bui, D., 2017. Rainfall-induced landslide susceptibility assessment at the Chongren area (China) using frequency ratio, certainty factor, and index of entropy. Geocarto. Int. 32 (2), 139–154. <https://doi.org/10.1080/10106049.2015.1130086>.

- Huang, F., Zhang, J., Zhou, C., Wang, Y., Huang, J., Zhu, L., 2020. A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. *Landslides* 17 (1), 217–229. <https://doi.org/10.1007/s10346-019-01274-9>.
- Jiang, Z., Fan, X., Subramanian, S.S., Yang, F., Tang, R., Xu, Q., Huang, R., 2021. Probabilistic rainfall thresholds for debris flows occurred after the Wenchuan earthquake using a Bayesian technique. *Eng. Geol.* 280, 105965 <https://doi.org/10.1016/j.enggeo.2020.105965>.
- Ju, L.Y., Xiao, T., He, J., Wang, H.J., Zhang, L.M., 2022. Predicting landslide runout paths using terrain matching-targeted machine learning. *Eng. Geol.* 311, 106902 <https://doi.org/10.1016/j.enggeo.2022.106902>.
- Kainthura, P., Sharma, N., 2022. Machine learning driven landslide susceptibility prediction for the Uttarkashi region of Uttarakhand in India. *Georisk* 16 (3), 570–583. <https://doi.org/10.1080/17499518.2021.1957484>.
- Kim, H., Lee, J.H., Park, H.J., Heo, J.H., 2021. Assessment of temporal probability for rainfall-induced landslides based on nonstationary extreme value analysis. *Eng. Geol.* 294, 106372 <https://doi.org/10.1016/j.enggeo.2021.106372>.
- Kirschbaum, D., Stanley, T., Zhou, Y., 2015. Spatial and temporal analysis of a global landslide catalog. *Geomorphology* 249, 4–15. <https://doi.org/10.1016/j.geomorph.2015.03.016>.
- Ko, F.W.Y., Lo, F.L.C., 2016. Rainfall-based landslide susceptibility analysis for natural terrain in Hong Kong – a direct stock-taking approach. *Eng. Geol.* 215, 95–107. <https://doi.org/10.1016/j.enggeo.2016.11.001>.
- Kocaman, S., Gokceoglu, C., 2019. A CitSci app for landslide data collection. *Landslides* 16 (3), 611–615. <https://doi.org/10.1007/s10346-018-1101-2>.
- 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>.
- Li, H.W.M., Lo, F.L.C., Wong, T.K.C., Cheung, R.W.M., 2022. Machine learning-powered rainfall-based landslide predictions in Hong Kong—an exploratory study. *Appl. Sci.* 12 (12), 6017. <https://doi.org/10.3390/app12126017>.
- Lin, Q., Wang, Y., 2018. Spatial and temporal analysis of a fatal landslide inventory in China from 1950 to 2016. *Landslides* 15 (12), 2357–2372. <https://doi.org/10.1007/s10346-018-1037-6>.
- Little, R.J., Rubin, D.B., 2020. *Statistical Analysis with Missing Data*, 3rd ed. John Wiley & Sons, Hoboken.
- Luo, J., Zhang, L., Yang, H., Wei, X., Liu, D., Xu, J., 2022. Probabilistic model calibration of spatial variability for a physically-based landslide susceptibility model. *Georisk* 16 (4), 728–745. <https://doi.org/10.1080/17499518.2021.1988986>.
- Medina, V., Hürlimann, M., Guo, Z., Lloret, A., Vaunat, J., 2021. Fast physically-based model for rainfall-induced landslide susceptibility assessment at regional scale. *Catena* 201, 105213. <https://doi.org/10.1016/j.catena.2021.105213>.
- Merghadi, A., Yunus, A.P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D.T., Avtari, R., Abderrahmane, B., 2020. Machine learning methods for landslide susceptibility studies: a comparative overview of algorithm performance. *Earth-Sci. Rev.* 207, 103225 <https://doi.org/10.1016/j.earscirev.2020.103225>.
- Meusburger, K., Alewell, C., 2008. Impacts of anthropogenic and environmental factors on the occurrence of shallow landslides in an alpine catchment (Urseren Valley, Switzerland). *Nat. Hazards Earth Syst. Sci.* 8 (3), 509–520. <https://doi.org/10.5194/nhess-8-509-2008>.
- Nadin, F., Kjekstad, O., Peduzzi, P., Herold, C., Jaedicke, C., 2006. Global landslide and avalanche hotspots. *Landslides* 3 (2), 159–173. <https://doi.org/10.1007/s10346-006-0036-1>.
- Nocedal, J., Wright, S., 2006. *Numerical Optimization*, 2nd edition. Springer Science + Business Media, New York.
- Pradhan, A.M.S., Lee, S.R., Kim, Y.T., 2019. A shallow slide prediction model combining rainfall threshold warnings and shallow slide susceptibility in Busan, Korea. *Landslides* 16 (3), 647–659. <https://doi.org/10.1007/s10346-018-1112-z>.
- Rahardjo, H., Nistor, M.M., Gofar, N., Satyanaga, A., Qin, X., Yee, S.I.C., 2020. Spatial distribution, variation and trend of five-day antecedent rainfall in Singapore. *Georisk* 14 (3), 177–191. <https://doi.org/10.1080/17499518.2019.1639196>.
- Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F., 2020. Dynamic path-dependent landslide susceptibility modelling. *Nat. Hazards Earth Syst. Sci.* 20 (1), 271–285. <https://doi.org/10.5194/nhess-20-271-2020>.
- Segoni, S., Piciullo, L., Gariano, S.L., 2018. A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides* 15 (8), 1483–1501. <https://doi.org/10.1007/s10346-018-0966-4>.
- Shen, P., Zhang, L., Chen, H., Fan, R., 2018. EDDA 2.0: integrated simulation of debris flow initiation and dynamics considering two initiation mechanisms. *Geosci. Model Dev.* 11 (7), 2841–2856. <https://doi.org/10.5194/gmd-11-2841-2018>.
- Steger, S., Brenning, A., Bell, R., Glade, T., 2017. The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. *Landslides* 14 (5), 1767–1781. <https://doi.org/10.1007/s10346-017-0820-0>.
- Su, C., Wang, B., Lv, Y., Zhang, M., Peng, D., Bate, B., Zhang, S., 2022. Improved landslide susceptibility mapping using unsupervised and supervised collaborative machine learning models. *Georisk* 1–19. <https://doi.org/10.1080/17499518.2022.208802>.
- Tehrani, F.S., Calvello, M., Liu, Z., Zhang, L., Lacasse, S., 2022. Machine learning and landslide studies: recent advances and applications. *Nat. Hazards* 114, 1197–1245. <https://doi.org/10.1007/s11069-022-05423-7>.
- Tonini, M., Pecoraro, G., Rommailler, K., Calvello, M., 2022. Spatio-temporal cluster analysis of recent Italian landslides. *Georisk* 16 (3), 536–554. <https://doi.org/10.1080/17499518.2020.1861634>.
- Van Den Eeckhaut, M., Hervás, J., Jaedicke, C., Malet, J.P., Montanarella, L., Nadim, F., 2012. Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. *Landslides* 9 (3), 357–369. <https://doi.org/10.1007/s10346-011-0299-z>.
- Van Westen, C.J., Rengers, N., Soeters, R., 2003. Use of geomorphological information in indirect landslide susceptibility assessment. *Nat. Hazards* 30 (3), 399–419. <https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e>.
- Wang, H., Zhang, L., Luo, H., He, J., Cheung, R.W.M., 2021. AI-powered landslide susceptibility assessment in Hong Kong. *Eng. Geol.* 288, 106103 <https://doi.org/10.1016/j.enggeo.2021.106103>.
- Xiao, T., Zhang, L.M., 2020. Evaluation of performance of engineered slopes under extreme rainstorms. In: *Geo-Congress 2020: Engineering, Monitoring, and Management of Geotechnical Infrastructure*, vol. 316. GSP, Reston, ASCE, pp. 737–743. <https://doi.org/10.1061/978084482797.072>.
- Xiao, T., Zhang, L.M., Cheung, R.W.M., Lacasse, S., 2022. Predicting spatio-temporal man-made slope failures induced by rainfall in Hong Kong using machine learning techniques. *Géotechnique* 1–17. <https://doi.org/10.1680/jgeot.21.00160>.
- Xing, X., Wu, C., Li, J., Li, X., Zhang, L., He, R., 2021. Susceptibility assessment for rainfall-induced landslides using a revised logistic regression method. *Nat. Hazards* 106 (1), 97–117. <https://doi.org/10.1007/s11069-020-04452-4>.
- Yalcin, A., Reis, S., Aydinoglu, A.C., Yomralioğlu, T., 2011. A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *Catena* 85 (3), 274–287. <https://doi.org/10.1016/j.catena.2011.01.014>.
- Yang, H.Q., Zhang, L., Gao, L., Phoon, K.K., Wei, X., 2022. On the importance of landslide management: Insights from a 32-year database of landslide consequences and rainfall in Hong Kong. *Eng. Geol.* 299, 106578 <https://doi.org/10.1016/j.enggeo.2022.106578>.
- Zhou, S.Y., Gao, L., Zhang, L.M., 2019. Predicting debris-flow clusters under extreme rainstorms: a case study on Hong Kong Island. *Bull. Eng. Geol. Environ.* 78 (8), 5775–5794. <https://doi.org/10.1007/s10064-019-01504-3>.