Machine learning-powered landslide forecasting: from initiation to mobility

Te Xiao & Li-Min Zhang
The Hong Kong University of Science and Technology, Hong Kong

ABSTRACT

Prompt prediction of landslide occurrence and movement in a future rainstorm is one of the most effective manners to cope with the increasing landslide risk in a changing climate. Despite the rapid development of many machine learning algorithms, most studies stay on landslide susceptibility mapping because of the challenging time-unknown and terrain-unmatched issues in landslide forecasting. This study proposes two novel machine learning strategies to predict the spatio-temporal distribution of landslides considering both initiation and mobility. Hong Kong is taken as an example to demonstrate the capacity of city-scale landslide forecasting using machine learning. The spatio-temporal evolution of both man-made slope failures and natural terrain landslides in a rainstorm can be well predicted using machine learning models, which can provide a powerful real-time decision-making tool for landslide early warning and risk management.

1 INTRODUCTION

Hong Kong has long struggled with landslides, because of its rainy climate, hilly terrain, and dense urban development (Cheung 2021). According to the records of the Geotechnical Engineering Office (GEO), over ten major storms have occurred every year and have triggered approximately 380 natural terrain landslides and 230 man-made slope failures annually. A severe rainstorm can impact a large area and trigger thousands of landslides in a short time, wreaking havoc on urban systems and causing enormous fatalities. For example, the record-breaking rainstorm on 6-9 June 2008 alone triggered about 2400 natural terrain landslides and 200 man-made slope failures over the territory of Hong Kong, most of which occurred within the critical four hours; the 18 June 1972 rainstorm triggered two catastrophic landslides on Po Shan Road and Sau Mau Ping that caused 138 deaths in a single day.

To manage landslide risk in a changing climate, in addition to structural landslide mitigation programmes that augment the capacity of slope safety system (e.g., Landslip Preventive Measures Programme (1977-2010) and Landslip Prevention and Mitigation Programme (2010-present) implemented by the GEO), non-engineering measures such as early warning and emergency management will be more effective in minimising the consequences of landslides. The GEO has operated a territory-wide landslide early warning system for 40 years (Kong et al. 2020), the latest version of which is based on statistical rainfall-landslide correlations (Xiao and Zhang 2020). The most critical question in landslide early warning concerned by decision-makers is "where and when would landslides occur amid a rainstorm?" This requires the prediction of spatio-temporal distribution of landslide occurrence. After initiation, landslides may travel certain distances downstream (e.g., hundreds of metres), threatening more people and properties along larger runout areas. For example, on 7 June 2008, massive landslide debris travelled over 600 m from its source to Yu Tung Road (Figure 1), resulting in the closure of westbound lanes for two months. The prediction of landslide mobility is thus another important issue in landslide forecasting.

Figure 1: The Yu Tung Road landslide on 7 June 2008

Physically-based methods (Baum et al., 2010; Shen et al. 2018; Kwan et al. 2021) can predict both the initiation and post-failure movement of landslides but are less suitable for real-time applications because of the high computational costs. By contrast, data-driven methods such as statistical correlation and machine learning (Finlay et al. 1999; Dai and Lee 2002; Merghadi et al. 2020; Wang et al. 2021) can make real-time predictions. Machine learning has emerged as a powerful alternative to traditional statistical methods, with the availability of larger amounts of landslide data and the advancement of computer technologies. Despite the rapid development of many algorithms, most machine learning-related landslide studies stay on susceptibility mapping that identifies landslide-prone areas from the spatial distribution of historical landslide records. The absence of time-related elements in landslide susceptibility mapping makes it fail to fully leverage the high efficiency of machine learning and fail to forecast landslide occurrence and movement in a future rainstorm.

This study will first overview the challenges in machine learning-powered landslide forecasting and then develop novel machine learning strategies to predict the spatio-temporal distribution of landslides considering both initiation and mobility. Hong Kong will be taken as an example to demonstrate the capacity of city-scale landslide forecasting using machine learning. The machine learning models can provide a powerful real-time decision-making tool for landslide early warning and risk management.

2 CHALLENGES IN LANDSLIDE FORECASTING

In the literature, spatio-temporal landslide occurrence prediction and post-landslide runout prediction are rarely investigated using machine learning techniques. This is mainly because (1) most landslide databases collected are not qualified for such purposes, particularly missing records of landslide times and paths; and (2) landslide initiation and mobility involve more complicated dynamic processes, thus requiring more advanced task-specific machine learning strategies than general algorithms. Specifically, the time-unknown issue and terrain-unmatched issue are the most challenging problems hindering the predictions of landslide initiation and mobility, respectively.

2.1 Time-unknown issue in landslide initiation forecasting

To develop a machine learning model that can predict landslide occurrence in a future rainstorm, it is vital to link historical landslides with their triggering rainstorms through the recorded failure times, known as storm-based data integration (Xiao et al. 2022). By this means, dynamic rainfall features of different rainstorms can be utilised for temporal prediction, in addition to various static slope and geo-information features as used in landslide susceptibility mapping.

Notably, the failure time is available for man-made slope failures but not for natural terrain landslides. Man-made slope failures in urban areas can be promptly reported by the public so that the failure time can be accurate to a day. By contrast, natural terrain landslides in remote areas are usually identified through post-event geological surveys or interpretation from remote sensing images over a certain period (e.g., once in a year or even less frequent), from which only the range of landslide time (i.e., within several possible rainstorms) can be determined. Conventional data processing cannot handle such time-unknown 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, another common method that assigns all time-unknown landslides to the largest storm in a given period 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 statistical rainfall-landslide correlations into storm-based ones. This adjustment approach has not been rigorously verified and may not be extensible to more complex scenarios.

2.2 Terrain-unmatched issue in landslide mobility forecasting

Regarding landslide mobility prediction, previous studies have established some region-specific statistical relationships between the fall height and travel distance, two key parameters describing a landslide runout path. As a matter of fact, both variables are unknown before the stop of landslide mass movement. With the unknown fall height, a trial-and-error terrain matching process is required to construct a 3-D runout path. Specifically, the longitudinal profile of a possible runout path starting from the landslide source (e.g., the steepest path) is first extracted from a 3-D digital terrain model (DTM), and all cells along the path are then visited one by one to determine where the landslide should terminate, namely the best location whose fall height and travel distance to source on the DTM satisfy the statistical correlation of the database. For statistical models, such a terrain matching process can be adopted for path prediction after the model training but is hardly involved during the model training. The inconsistency in terrain matching may reduce the accuracy of landslide runout path prediction (Ju et al. 2022), despite a strong correlation between the fall height and travel distance.

3 MACHINE LEARNING-POWERED LANDSLIDE FORECASTING

3.1 Comprehensive databases for Hong Kong

Hong Kong is a society rich in landslide-related data, which provides an ideal opportunity to release the potential of machine learning in landslide forecasting. Fundamental databases for this task should cover rainstorms (Figure 2(a)), slopes (Figure 2(b)), and landslides (Figure 2(c)). In the period of 1984–2017, 419 major rainstorms are identified, and the hourly rainfall amounts of 50 Hong Kong Observatory (HKO) rain gauges and 91 GEO rain gauges are collected. The slope information system managed by the GEO had registered more than 60,000 man-made slopes of cut slopes, fill slopes, and retaining walls.

Past landslides in Hong Kong have been compiled into two databases. One is the Enhanced Natural Terrain Landslide Inventory (ENTLI) (https://www.geomap.cedd.gov.hk/GEOOpenData/eng/ENTLI.aspx), 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 ENTLI takes points and polylines to represent the landslide sources and paths, respectively, and records four key geometric features, namely travel distance, fall height, and length and width of the landslide source.

The catalogue of landslide incidents (https://www.geomap.cedd.gov.hk/GEOOpenData/eng/Incident.aspx) is another major landslide database, which records a few reported natural terrain landslides and about 8000 man-made slope failures in 1984–2017. Prompt reporting from the public makes the failure time accurate to a day in most incidents. The tabular catalogue of landslide incidents contains the time, location, type, scale, and consequences of landslides.

In addition to rainfall, slope, and landslide databases, geographic databases of terrain, geology, land cover, and infrastructure are also utilised in machine learning in this study. Given such rich data, the abovementioned time-unknown issue and terrain-unmatched issue can be properly addressed through new machine learning strategies developed as follows.

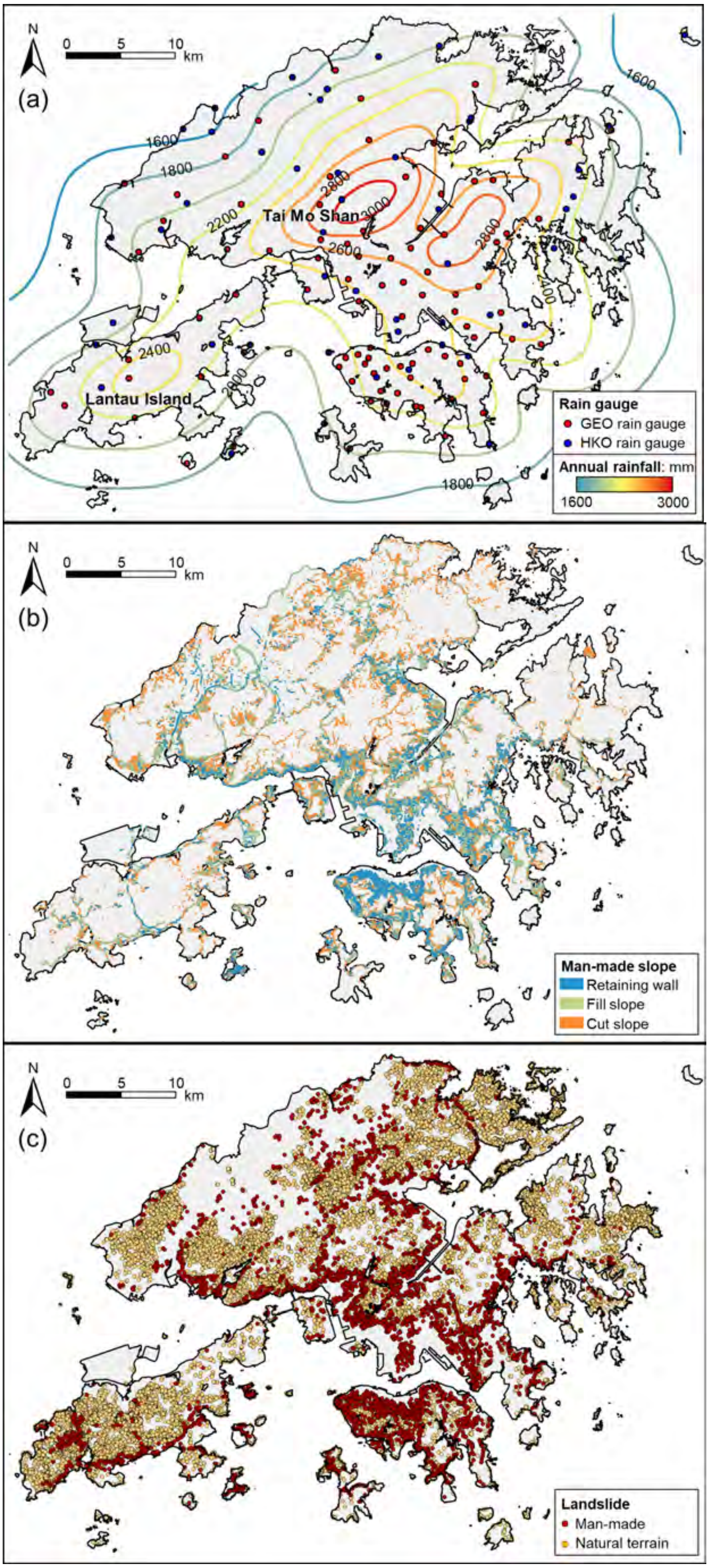

Figure 2: Comprehensive databases for Hong Kong: (a) rainstorm; (b) slope; (c) landslide

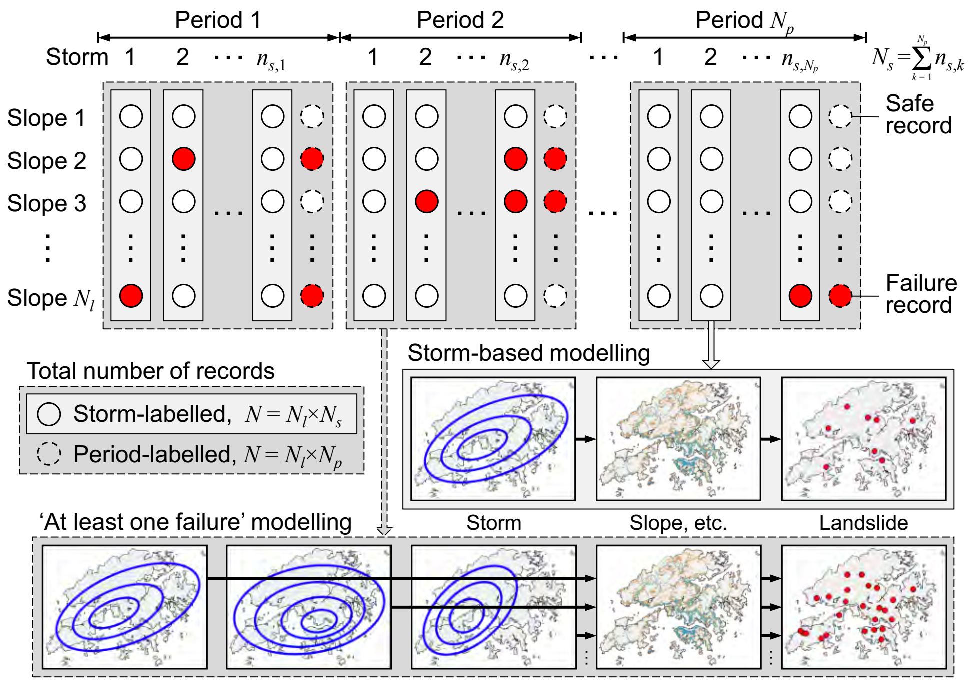

Figure 3: Probabilistic modelling strategy for landslide initiation forecasting

3.2 *Probabilistic modelling strategy for landslide initiation forecasting*

If landslide times are known (i.e., storm-labelled landslides as shown in Figure 3), a storm-based landslide forecasting model can be directly established using any machine learning algorithm (Xiao et al. 2022), similar to the procedure of landslide susceptibility mapping but involving dynamic rainfall features such as the maximum rolling 24-h rainfall amount to achieve temporal prediction.

If landslide times are unknown but labelled over a consistent period (e.g., one year for year-labelled landslides in ENTLI), it is conservative to claim that, for a particular landslide, the slope has at least one failure triggered by storms in this period. Therefore, N_l slopes after N_p-period observations produce a total of $N = N_l \times N_p$ records (Figure 3). 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 as (Xiao and Zhang 2023a):

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

where p_s = failure probability of a slope in the sth storm, estimated by a storm-based landslide forecasting model. Equation (1) is referred to as probabilistic modelling strategy of 'at least one failure' for time-unknown landslides. Consequently, the cost function $J(\boldsymbol{\theta})$ of the machine learning model given a set of model parameters $\boldsymbol{\theta}$ can be written as:

$$J(\boldsymbol{\theta}) = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i \ln P_i + (1 - y_i) \ln (1 - P_i) \right] \tag{2}$$

where y_i = period-based observation of the ith record: $y_i = 0$ for a non-landslide record and 1 for a landslide record; and P_i = predicted failure probability of the ith record using Equation (1).

Equation (2) is similar to the cost function used in storm-based forecasting for time-known landslides, but the landslide observation 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 are 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, Equation (2) still computes $N_l \times N_s$ failure probabilities (where N_s is the total

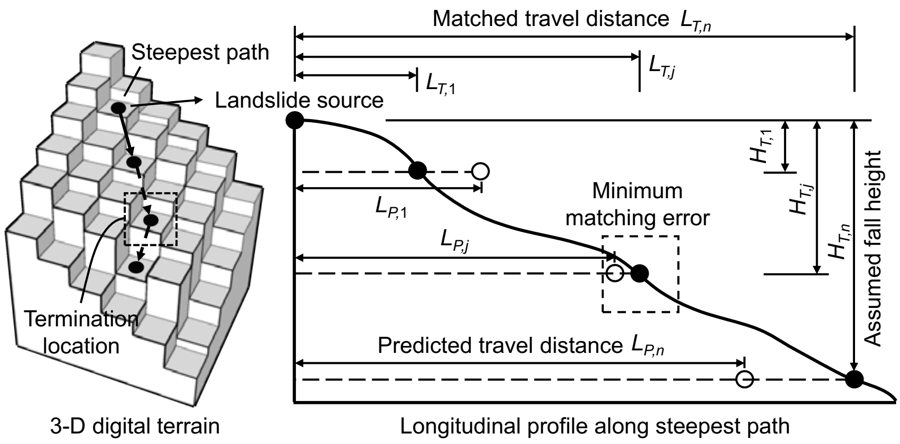

Figure 4: Trial-and-error terrain matching for landslide runout path prediction

number of storms in all periods), consistent with that in a storm-labelled scenario. This enables storm-based landslide forecasting using time-unknown landslide data.

The realistic landslide database is more often a mixture of both time-known and time-unknown landslides, and the observation period may vary from one slope to another. In such a case, an integrated machine learning model can be established using the proposed 'at least one failure' strategy, and the cost function is a weighted combination of time-known and time-unknown landslides (Xiao and Zhang 2023a).

During the landslide initiation forecasting, it is also possible to simultaneously incorporate the prediction of landslide scale (Xiao et al. 2022), by re-formulating the binary classification problem (i.e., non-landslide and landslide) as a multiclass classification problem according to various landslide scales (e.g., non-landslide, and very minor, minor, major, and very major landslides). The landslide scale will have a great impact on the subsequent prediction of landslide mobility.

3.3 Terrain matching strategy for landslide mobility forecasting

Traditionally, a statistical model (e.g., multivariate linear regression) only predicts the landslide travel distance; hence, a trial-and-error terrain matching process is required to convert the distance into a path. As shown in Figure 4, it requires a 3-D DTM and a landslide runout analysis model that predicts a travel distance (L) from a given fall height (H) and other factors. A possible runout path starting from the source is identified first on the DTM by assuming that the landslide moves along the steepest path to an adjacent position. If the landslide stops at the jth cell along the path (j = 1, 2, …, n), the actual fall height and travel distance on the DTM from the landslide source to the termination location are $H_{T,j}$ and $L_{T,j}$, respectively, while the runout analysis model predicts another travel distance $L_{P,j}$ from the given $H_{T,j}$. The two distances (i.e., $L_{T,j}$ and $L_{P,j}$) may not be identical. After going through all cells on the path, the cell with a minimum matching error between $L_{T,j}$ and $L_{P,j}$ can be taken as the termination location, and the runout path is accordingly determined. Note that the targets of model training and path prediction are inconsistent in a statistical model. For model training, the target is to fit the predicted travel distance (L_P) to the one recorded in the landslide database (L), while it is to match the predicted travel distance to the one derived from the DTM (L_T) in path prediction.

A consistent terrain matching strategy in both model training and prediction is proposed for machine learning-based landslide runout path prediction (Ju et al. 2022), as shown in Figure 5. Unlike conventional machine learning models, an additional step for determining the most probable termination location is inserted between the forward runout prediction and backward parameter estimation. The explanatory variables of each cell along the possible path are fed separately into the machine learning model to predict multiple travel distances $L_{P,j}$ (i.e., Step II). Afterwards, the cell k with a minimum matching error between $L_{P,j}$ and the actual distance $L_{T,j}$ on the DTM is regarded as a temporary termination cell (i.e., Step III), and the actual distance $L_{T,k}$ is then fed back into the machine learning model to estimate model parameters (i.e., Step IV). These procedures are iteratively repeated until the optimal model parameters $\boldsymbol{\theta}$ are obtained. The cost function to be minimised can be written as:

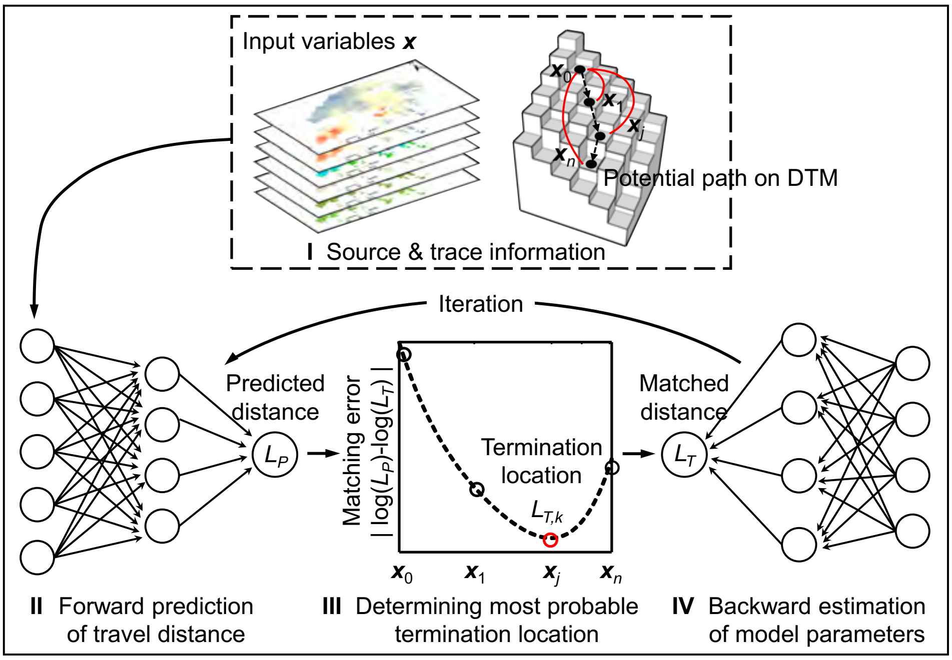

Figure 5: Terrain matching strategy for landslide mobility forecasting

$$J(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}\left[\log\left(L_{T,k,i}\right) - \log\left(L_i\right)\right]^2 \tag{3}$$

where $L_{T,k,i}$ = the most probable travel distance at location k on the DTM for the ith landslide and the termination location k can be determined according to the fitness of terrain as:

$$k = \underset{j=1,2,\ldots,n_i}{\arg\min}\left|\log\left(L_{P,j,i}\right) - \log\left(L_{T,j,i}\right)\right| \tag{4}$$

where $L_{P,j,i}$ = predicted travel distance at the jth cell on the ith landslide path.

Combining Equations (4) and (3) makes the predicted travel distance (L_P) first fit the one on the DTM (L_T) and then the one recorded in the landslide database (L). Such a pseudo bi-objective optimization makes the predicted landslide path more comparable with the terrain reality. Another advantage of the terrain matching iteration is that: it can easily consider the variable trace features from the landslide source to the jth cell when visiting all possible termination locations.

It should be highlighted that the proposed machine learning strategies to address the time-unknown issue in initiation forecasting and terrain-unmatched issue in mobility forecasting can be flexibly applied with various basic machine learning algorithms, such as logistic regression, neural network, or even deep learning techniques, as they only need the updating of cost functions. The validation and comparison of the proposed machine learning strategies against historical landslide data and statistical models can be referred to Xiao and Zhang (2023a) and Ju et al. (2022). The following section will take Hong Kong as an example to demonstrate the capacity of city-scale landslide forecasting using machine learning.

4 CASE STUDIES FOR HONG KONG

4.1 Spatio-temporal landslide forecasting

Applying the probabilistic modelling strategy, regardless of whether the landslide time is known or not, two landslide initiation forecasting models can be developed for the 60,000 registered man-made slopes (Xiao et al. 2022) and 26 million natural terrain hillslope cells (5 m × 5 m) (Xiao and Zhang 2023b), respectively. Logistic regression is adopted as the basic machine learning classifier in both models.

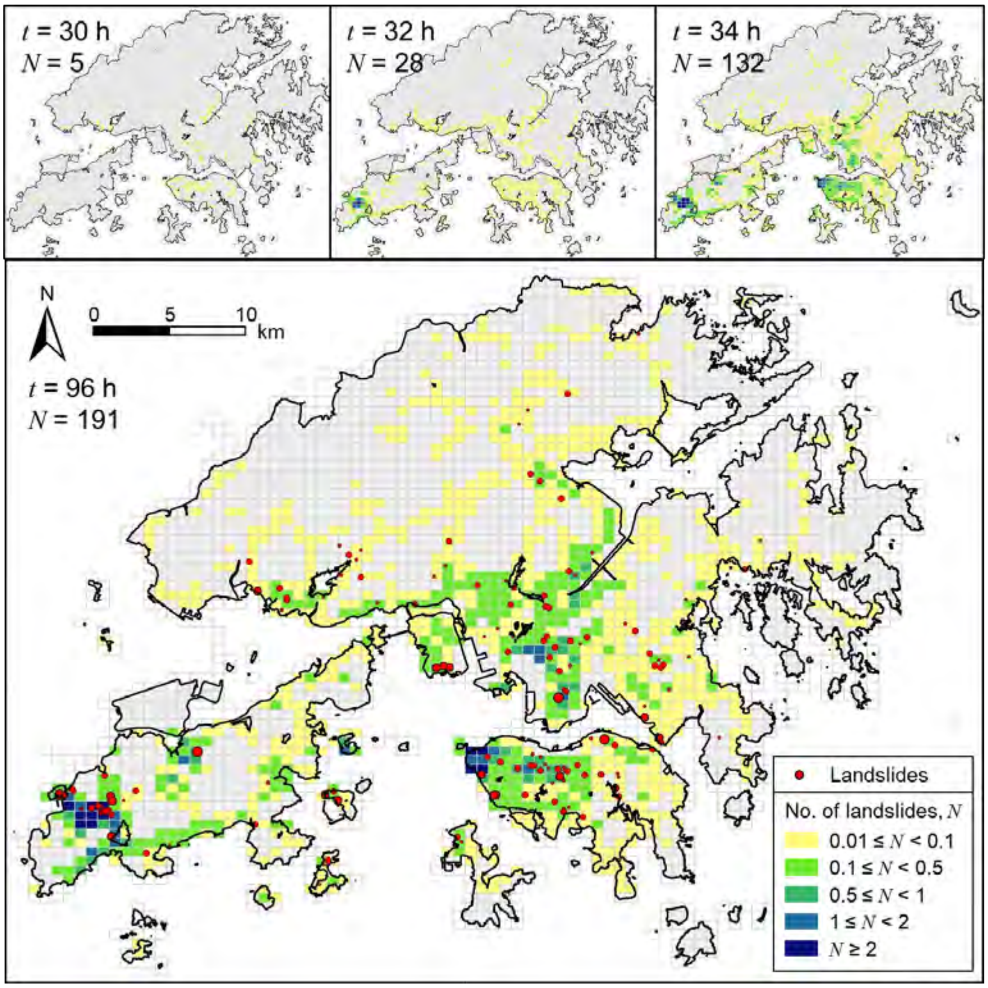

Figure 6: Predicted evolution of man-made slope failures during the 6-9 June 2008 storm

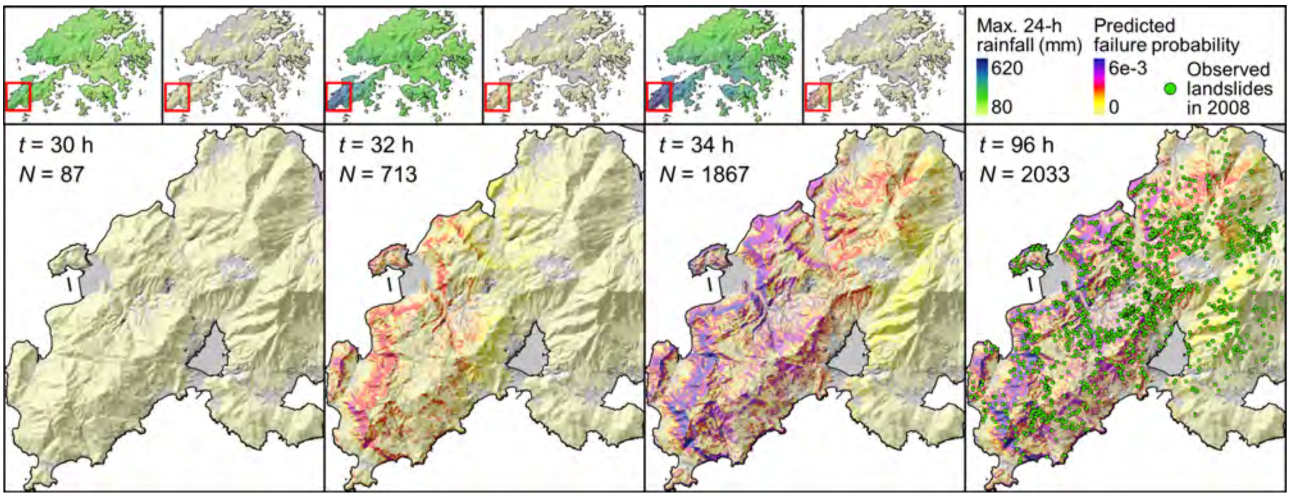

Figure 7: Predicted evolution of natural terrain landslides during the 6-9 June 2008 storm

The 6-9 June 2008 storm is one of the most severe rainstorms in the history of Hong Kong, with 1-h, 4-h, and 24-h rainfall amounts all falling in the top five rainstorms ever recorded. The storm triggered 162 confirmed failures of registered man-made slopes and about 2400 natural terrain landslides (estimated). The

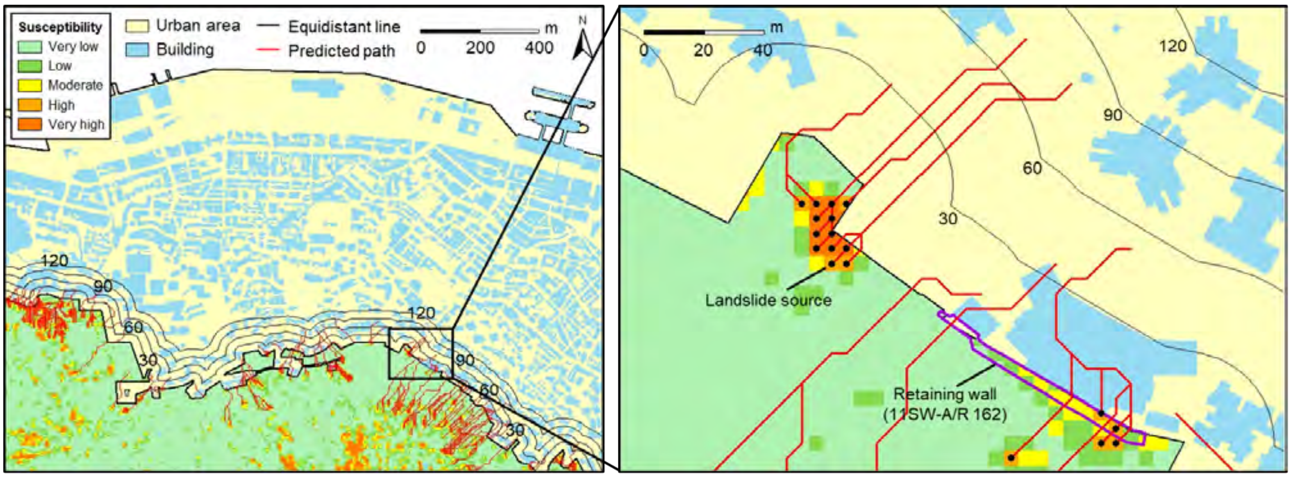

Figure 8: Predicted landslide-affected areas in the Mid-Levels

spatio-temporal evolutions of man-made slope failures (Figure 6) and natural terrain landslides (Figure 7) during the 6-9 June 2008 storm can be predicted using the two machine learning model with nearly real-time efficiency. For man-made slopes, hardly any landslides occur before time t = 30 h; the western Lantau Island is the first area to witness many landslides at t = 32 h; afterwards, there is a rapid outbreak of landslides on the north-western Hong Kong Island and western Kowloon. The predicted number of natural terrain landslides experiences a similar temporal escalation, from less than 100 at t = 30 h to more than 1800 at t = 34 h, but almost all concentrate on western Lantau Island in space. Both the predicted numbers of landslides and landslide-prone areas agree well with the observations.

4.2 Identification of landslide-affected urban areas

Applying the terrain matching strategy, the performance of the machine learning-powered landslide mobility forecasting model can be significantly improved from less than 0.4 to higher than 0.7, even though a very simple neural network is adopted (Ju et al. 2022). The proposed machine learning model can not only provide rapid regional runout path predictions like a statistical model does, but also reasonably incorporate complex geographic characteristics along landslide traces and 3-D terrain reality like in a 3-D numerical simulation.

With the developed prediction model of landslide runout paths, it is possible to identify potential landslide-affected urban areas and high-risk landslide-bearing elements. Consider the Mid-Levels at the foot of Victoria Peak on Hong Kong Island as an example (Figure 8). The Mid-Levels is a landslide-prone area with various buildings densely packed near the steep natural hillsides. One of the most severe rainstorms hitting the Mid-Levels in history is the storm on 16-21 August 2005, with the maximum rolling 4-h and 24-h rainfall amounts being 171 mm and 567.5 mm, respectively. Integrated with landslide susceptibility mapping results, 2238 landslide travel paths can be predicted in the investigated area, among which 1268 landslides rush into the urban area, as shown in Figure 8. About 72% of these landslides stop within 30 m to the mountain foot, and only a few travel a long distance (e.g., 0.5% over 90 m to the mountain foot). The building at the lower right corner of Figure 8 is at the highest risk of being attacked by several potential landslides and demands necessary slope stabilization measures. In fact, an 8 m high concrete retaining wall (11SW-A/R 162) has been constructed between the building and the mountain for safety reasons.

4.3 Future application to prompt quantitative risk assessment

The ultimate application target of the machine learning-powered landslide forecasting models will be the prompt quantitative risk assessment (He et al. 2023), which requires real-time predictions of both landslide initiation (i.e., occurrence probability) and mobility (i.e., spatial impact probability) under a full probabilistic framework (Figure 9).

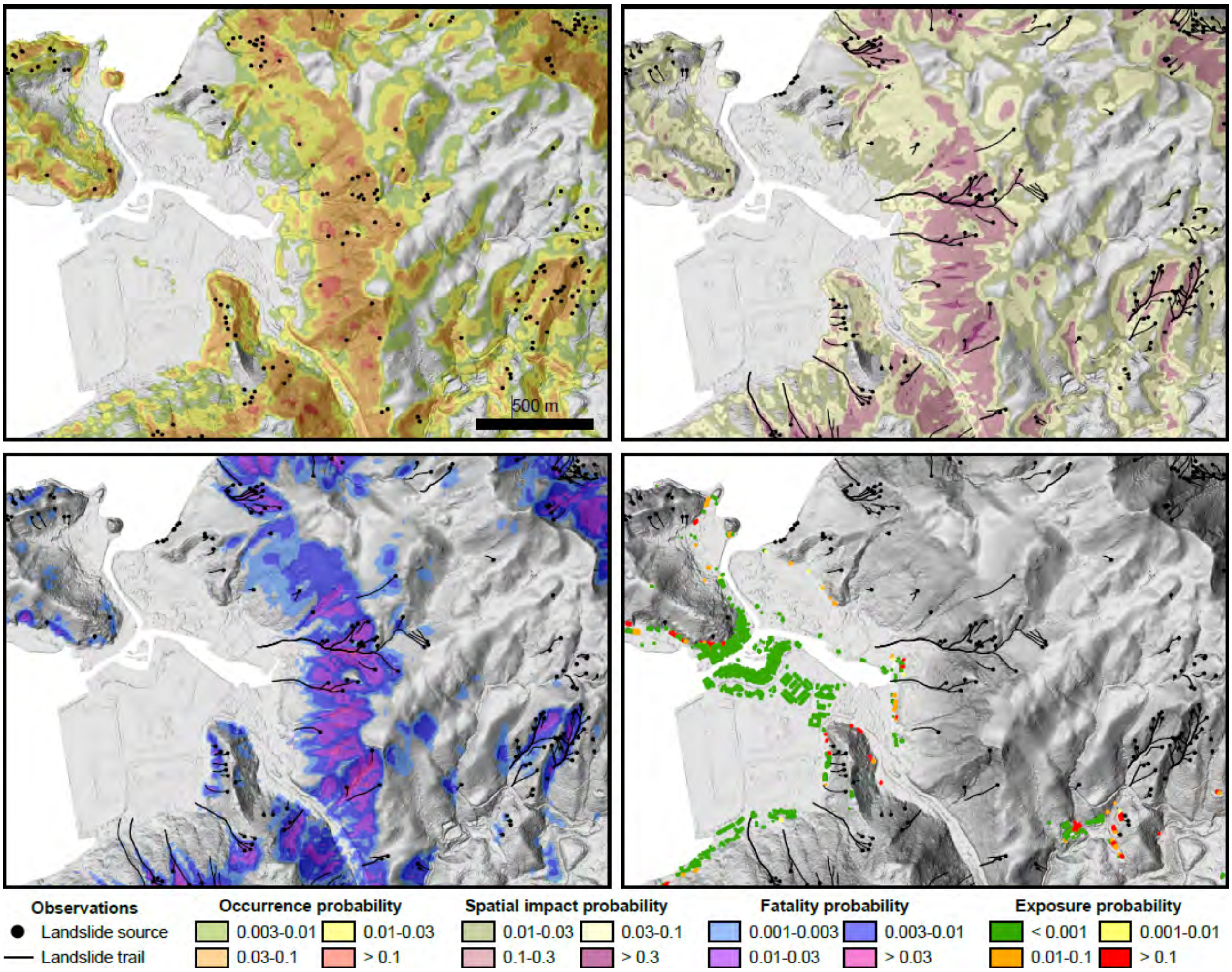

Figure 9: Example of quantitative landslide risk assessment

The current statistical version only assesses the risk of natural terrain landslides. Considering that man-made slope failures occurring in densely populated urban regions (e.g., Hong Kong Island and Kowloon) will pose much higher risk than natural terrain landslides occurring in sparsely populated remote regions (e.g., Lantau Island), it is crucial to assess the risk of man-made slope failures simultaneously. As demonstrated in the previous two applications, machine learning models have already achieved the initiation prediction of both natural terrain landslides and man-made slope failures. Future work should focus on the mobility prediction of man-made slope failures. A possible solution is to discretize man-made slopes into the same cells as the natural terrain does. By this means, the post-failure movement of man-made slopes can be modelled in a consistent probabilistic manner.

The GEO is operating two landslide early warning systems for man-made slope failures and natural terrain landslides separately. Due to the difference in safety levels and covering areas of these two types of landslides, different quantity-based warning criteria are adopted, namely 15 for man-made slope failures and [500, 1000, 2000] for natural terrain landslides. The landslide risk as an integration of both occurrence probability and consequence would be a better indicator than the landslide quantity to unify man-made slope failures and natural terrain landslides. The next-generation risk-informed landslide early warning system will benefit a lot from machine learning-powered landslide forecasting.

5 CONCLUSIONS

Machine learning-related landslide studies stay on landslide susceptibility mapping for decades. Moving towards the prediction of landslide occurrence and movement, the time-unknown issue and terrain-unmatched issue should be addressed primarily. For such a purpose, this study proposes two novel machine learning strategies of probabilistic modelling and terrain matching for forecasting landslide initiation and mobility, respectively. It is found that even simple machine learning algorithms can have significant improvements than conventional statistical methods, as long as well-designed task-specific machine learning strategies can be developed.

Hong Kong is taken as an example to demonstrate the capacity of city-scale landslide forecasting using machine learning. The spatio-temporal evolution of both man-made slope failures and natural terrain landslides in a rainstorm can be well predicted using machine learning models, which can provide a powerful real-time decision-making tool for landslide early warning and risk management.

ACKNOWLEDGEMENTS

This work was supported by the Research Grants Council of the Hong Kong SAR (Nos. 16205719, 16203720, N_HKUST620/20, and AoE/E-603/18). The authors would like to thank the Geotechnical Engineering Office of the Civil Engineering and Development Department, the Government of the Hong Kong SAR, for providing the landslide data.

REFERENCES

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. *Journal of Geophysical Research: Earth Surface*, 115: F03013.

Cheung, R.W.M. 2021. Landslide risk management in Hong Kong. *Landslides*, 18(10): 3457-3473.

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.

Finlay, P.J., Mostyn, G.R., & Fell, R. 1999. Landslide risk assessment: prediction of travel distance. *Canadian Geotechnical Journal,* 36(3): 556-562.

He, J., Zhang, L., Xiao, T., Wang, H., & Luo, H. 2023. Prompt quantitative risk assessment for rain-induced landslides. *Journal of Geotechnical and Geoenvironmental Engineering*, 149(5): 04023023.

Ju, L.Y., Xiao, T., He, J., Wang, H.J., & Zhang, L.M. 2022. Predicting landslide runout paths using terrain matching-targeted machine learning. *Engineering Geology*, 311: 106902.

Ko, F.W., & Lo, F.L. 2016. Rainfall-based landslide susceptibility analysis for natural terrain in Hong Kong-A direct stock-taking approach. *Engineering Geology*, 215: 95-107.

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.

Kwan, J.S.H., Sze, E.H.Y., Lam, C., Law, R.P.H., & Koo, R.C.H. 2021. Development and applications of debris mobility models in Hong Kong. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, 174(5): 593-610.

Merghadi, A., Yunus, A.P., Dou, J., et al. 2020. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, 207: 103225.

Shen, P., Zhang, L., Chen, H., & Fan, R. 2018. EDDA 2.0: integrated simulation of debris flow initiation and dynamics considering two initiation mechanisms. *Geoscientific Model Development*, 11(7): 2841-2856.

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.

Xiao, T., & Zhang, L. 2020. Evaluation of performance of engineered slopes under extreme rainstorms. In *Geo-Congress 2020: Engineering, Monitoring, and Management of Geotechnical Infrastructure*. pp. 737-743. Reston, United States: ASCE.

Xiao, T., & Zhang, L.M. 2023a. Data-driven landslide forecasting: Methods, data completeness, and real-time warning. *Engineering Geology*, 317: 107068.

Xiao, T., & Zhang, L.M. 2023b. Storm-based forecasting of natural terrain landslides. In *Geo-Risk 2023: Advances in Theory and Innovation in Practice*. pp. 1-10. Reston, United States: ASCE.

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. DOI: 10.1680/jgeot.21.00160