K. T. Chau · J. E. Chan

Regional bias of landslide data in generating susceptibility maps using logistic regression: Case of Hong Kong Island

Abstract On the basis of 1,834 landslide data for Hong Kong Island (HKI), landslide susceptibility maps were generated using logistic regression and GIS. Regional bias of the landslide inventory is examined by dividing the whole HKI into a southern and a northern region, separated by an east-west trending water divide. It was found that the susceptibility map of southern HKI generated by using the southern data differs significantly from that generated by using northern data, and similar conclusion can be drawn for the northern HKI. Therefore, a susceptibility map of HKI was established based on regional data analysis, and it was found to reflect closely the spatial distributions of historical landslides. Elevation appears to be the most dominant factor in controlling landslide occurrence, and this probably reflects that human developments are concentrated at certain elevations on the island. Classification plot, goodness of fit, and occurrence ratio were used to examine the reliability of the proposed susceptibility map. The size of landslide susceptible zones varies depending on the data sets used, thus this demonstrates that the historical landslide data may be biased and affected by human activities and geological settings on a regional basis. Therefore, indiscriminate use of regional-biased data should be avoided.

Keywords Landslide data · Inventory · GIS · Regional bias · Hong Kong · China

Introduction

Different methods for landslide susceptibility mapping

Susceptibility map is found very useful in estimating, managing and mitigating landslide hazard for a region (e.g. Corominas and Santacana 2003; Sassa et al. 2004). The terms of landslide hazard and landslide susceptibility have been used in the literature in an interchangeable manner (e.g. Saha et al. 2005), but hazard normally include magnitude and frequency whereas susceptibility usually does not (Corominas et al. 2003). In this paper, we will mainly use the term susceptibility instead of hazard. Many review articles have addressed the issue of landslide susceptibility, risk analysis and management, including Leroi (1997), Hansen (1984), Fell and Hartford (1997), Einstein (1988, 1997), Varnes (1984), and Hungr (1997). When landslide susceptibility is estimated, historical landslide records were usually correlated to the local geology, lithology, structure, geomorphology, hydrologic conditions, vegetation, and climate. Although statistical means may be used for such correlation analysis, the most appropriate and popular choice appears to be the logistic regression (e.g. Marquinez et al. 2003; Menendez-Duarte et al. 2003; Rowbotham and Dudycha 1998; Dai and Lee 2001, 2002, 2003; Dai et al. 2001; Myster et al. 1997; Lee and Min 2001; Chau et al. 2004b; Ohlmacher and Davis 2003; Yesilnacar and Topal 2005; Wang and Sassa 2005; Lee 2004, 2005a, b; Ayalew and Yamagishi 2005). For example, Rowlotham and Dudycha (1998)

concluded that logistic regression was the most useful method for landslide. Ideally, a reliable landslide susceptibility map should carry appropriate weights from historical landslide events, from geomorphological analysis, and from mechanics analysis of slides (e.g. Chau and Lo 2004). Since all three aspects of susceptibility analysis involve the handling and interpreting a large amount of factual, geological and simulated data, the use of computer or information technology is essential. Geographical Information System (GIS) appears to be an appropriate choice (Coppock 1995). Some of these GIS-based susceptibility analyses focus on earthquake-induced landslides (e.g. Luzi et al. 2000; Miles and Keefer 1999; Refice and Capolongo 2002), and some focus on rockfall susceptibility (e.g. Cancelli and Crosta 1994).

Aims and objectives of this study

As mentioned in Chau et al. (2004a), GIS technology has been used in analyzing landslide data in many different parts of the world, including both developed and developing countries. However, none of the previous studies examines the possibility of regional bias of the landslide data. In other words, the landslide dependency on slope angle, geology, and slope aspect may change from one region to another. When a region consists mainly of south-facing and north-facing slopes, the amount of sunshine, rainfall and wind received by the north-facing and south-facing slopes may not be the same. This difference may play a key role in the process of erosion and weathering in such a way that the possibility of landslide occurrence may also vary. Therefore, it seems necessary and reasonable to first examine whether the local landslide inventory may reflect such differences, otherwise the landslide data analysis may not yield reliable predictions. However, to the best of our knowledge, there is no comprehensive study on the probable regional dependence of landslide data within an area. The only related study is on the size of landslide buffer zones around roads by Larsen and Parks (1997).

In the case of Hong Kong, GIS-assisted landslide hazard analysis has only been proposed for Lantau Island, where the new International Airport is located (Dai et al. 2001, 2002; Dai and Lee 2001, 2002; Lee et al. 2001). However, most of the population of Hong Kong does not locate at Lantau Island, but instead on the Hong Kong Island. Therefore, both rockfall (Chau et al. 2003, 2004b) and landslide (Chau et al. 2004a) hazards for Hong Kong Island have been considered. However, as mentioned in the last paragraph, the possibilities of regional bias of the data have not been considered.

Therefore, the main objective of this study is to investigate whether there is a regional bias of the landslide data in Hong Kong. More specifically, the northern and southern parts of the Hong Kong Island will be analyzed separately using logistic regression. We should emphasize that the landslide data of Hong Kong Island has not be analyzed before using logistic regression. The related study by Chau et al. (2004a) used a simple statistical weighting approach similar to that of Ayalew et al. (2004). As shown in Fig. 1, the northern Hong

Fig. 1 An aerial photograph showing the east-west trending water divide line for Hong Kong Island. The dotted line indicates the water divide for the north- and south-facing regions. In the lower left aerial photo of the whole Hong Kong, the arrow and solid line show the view of direction of the larger photograph and water divide line respectively
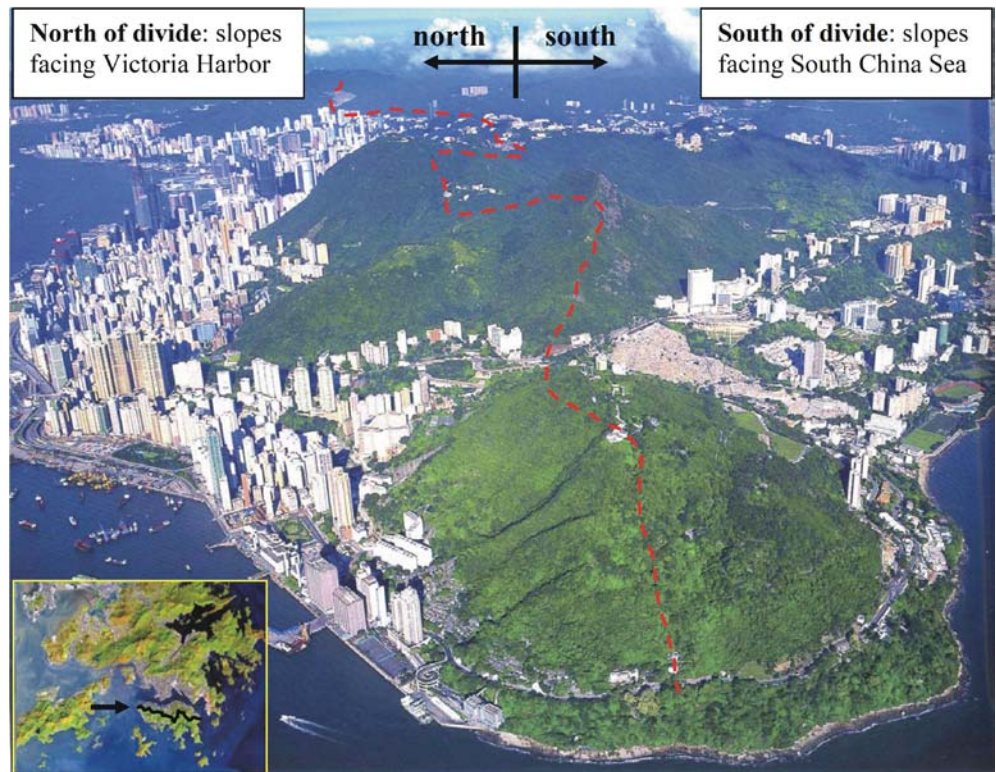
Kong Island mainly consists of north-facing slopes while the southern part mainly consists of south-facing slopes. The results of these regional data analyses will be compared to that obtained by using all landslide data of the whole Hong Kong Island. It is well known that Victoria Harbor (which is located just north of the Hong Kong Island) is sheltered from bad weather especially during typhoon season. As shown in Fig. 1 the urban developed in the northern Hong Kong Island is much more rapid. Geologically, the southern Hong Kong Island mainly consists of tuff while the northern Hong Kong Island consists of both tuff and granite. In addition, the slopes south of the water divide are more gentle whereas slopes north of the divide are generally steeper. Therefore, all these differences may lead to regional bias in the landslide occurrence.

In the next section, a brief summary of the geological condition and the landslide inventory for Hong Kong Island will first be given; in the section Logistic Regression Analysis for the landslide data, logistic regression analysis is introduced and applied to the whole Hong Kong Island using all, southern portion and northern portion of the landslide data respectively; finally, the Discussions and conclusion section concludes the paper.

Hong Kong Island and landslides inventory

Geological and weather condition of Hong Kong Island

The geological settings and the typical climatic situations of the Hong Kong Island have been discussed by Chau et al. (2004a) and therefore will not be repeated here. Only a brief summary is given here. Hong Kong Island is the second largest island in the territory of Hong Kong, but only comprises 7% of the total 1,098 km² land area of Hong Kong. The terrain of Hong Kong Island is hilly with steep slopes exceeding 30°, covered by superficial deposits. In most places the average thickness of the deposits is about 2 m, but in some places up

to 30 m (Fyfe et al. 2000). The highest point is the Victoria Peak at 554 m, and to the east of it there are 439 m Mount Cameron and 507 m Mount Parker. The island is underlain mainly by volcanic rock and intrusive granite. Hong Kong's climate is subtropical, with a winter temperature of 10°C to summer temperature of exceeding 31°C. Foggy weather is expected in spring, hot and humid weather in summer. When a typhoon comes close to Hong Kong, rain can become heavy; and landslides are very common in typhoon season from May to September. The mean annual rainfall ranges from around 1,300 mm along the coast to more than 3,000 mm on mountains. About 80% of the rain falls between May and September, with August being the wettest month.

Landslide data used in this study

In this section, landslide inventory used for the present susceptibility analysis is described briefly. They are mainly extracted from the reports on "Hong Kong rainfall and landslides" which are published annually by the Geotechnical Engineering Office (GEO) since 1984 and from a report compiling all landslide consequences since 1948 (GEO 1996). The current landslide inventory was supplied by the GEO from 1982 to 2002, with a total of 7,726 landslides in the period of 21 years and 1,834 of these from Hong Kong Island (23.7% of the total data). Consider the fact that only 7% of the land of Hong Kong is from Hong Kong Island, the rate of landslide occurrence on Hong Kong Island is considerably higher than the rest of Hong Kong. This is partially due to the fact that Hong Kong Island is the most populated area in Hong Kong and many of its roads, buildings, and facilities are built on steep slopes. The chance of not reporting a sizable landslide event on Hong Kong Island is slim, whereas many landslides (especially those on natural slopes in rural areas or country parks) in the New Territories of Hong Kong had never been reported if they did not disturb normal human activities.
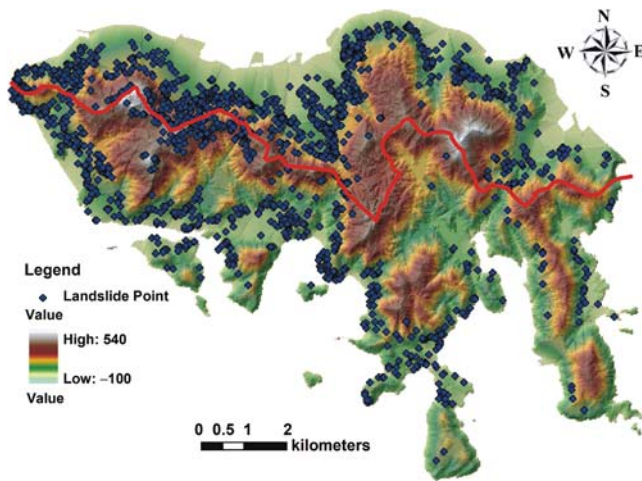
Fig. 2 The locations of 1,834 landslide events on Hong Kong Island from 1982 to 2002 are shown by the solid rhombus points, together with the background terrain relief map of TIN (Triangulated Irregular Network) format generated by contour map using ArcMap. The red line is the west-east trending divide line shown in Fig. 1; and the scale bar is in km

Table 1 The distribution of elevation of Hong Kong Island and the distribution of 1,834 landslide events among six different elevation ranges

Elevation ranges (m)	Elevation distribution (%)	Landslide distribution (%)
0–50	28	31
50–100	17	32
100–150	14	16
150–200	12	7
200–250	9	3
≥250	20	11

Table 2 The distribution of slope angle of Hong Kong Island and the distribution of 1,834 landslide events among eight different slope angle ranges

Elevation ranges (°)	Slope angle distribution (%)	Landslide distribution (%)
0–10	28	28
10–15	7	12
15–20	8	10
20–25	12	11
25–30	16	13
30–35	14	11
35–40	8	8
≥40	7	7

To examine regional bias, three landslide susceptibility analyses have been conducted: (1) a susceptibility map of using all 1,834 data; (2) a susceptibility map for southern Hong Kong using the landslide data and geological information of the south of the water divide shown in Figs. 1 and 2 (a total of 719 landslides); and (3) a susceptibility map for northern Hong Kong using the landslide data and geological information north of the water divide shown in Figs. 1 and 2 (a total of 1,115 landslides). Figure 2 shows the locations of all 1,834 landslide points together with the east-west trending divide line and a background terrain relief map in TIN (Triangular Irregular Network) format generated by using ArcGIS.

Elevation distribution of Hong Kong Island and landslides

Some statistics of the landslide data will be briefly discussed. Using the digital map of Hong Kong (scale of 1:5,000) and the "zonal statistics" function under "Spatial Analyst" in ArcGIS Version 8.2 (for a grid of 30 m×30 m), we can extract the information of the slope angle, slope aspect orientation, elevation and geology type of each landslide event. In particular, ArcGIS Version 8.2 is an updated version of ArcView 3.2 by ESRI that allows for more functional options. In this study, we mainly use ArcMap with "Spatial Analyst" and "3D-Analyst" options and use ArcCatalog available in ArcGIS.

In terms of the elevation, the percentages of landslides occurred within the ranges of elevation of 0–50 m, 50–100 m, 100–150 m, 150–200 m, 200–250 m and >250 m and the natural distributions of elevation of Hong Kong Island are shown in Table 1. The data shown in Table 1 is obtained by using the "zonal statistics" function. Table 1 shows that landslides are more likely to occur at an elevation of 0–100 m. It seems not unreasonable because, as shown in Fig. 1, most of the population is located on the reclaimed land or on the lower part of the slopes; therefore, most of the cut slopes are concentrated in the 0–100 m elevation range. In other words, the elevation dependence implicitly incorporates effect of roads and residential development (or human activities) at low elevation into the present analysis. As shown by Larsen and Parks (1997), the relation between human activities and landslides can also be studied using buffer zones of varying sizes. Note also that since landslide locations are reported at where landslides stop, but not where they are initiated. Therefore, the elevation of landslide occurrence in the inventory is likely of bias toward the lower ends.

Slope angle distribution of Hong Kong Island and landslides

In terms of the slope angle, the percentages of landslides occurred within the ranges of 0–10°, 10–15°, 15–20°, 20–25°, 25–30°, 30–35°, 35–40° and ≥40° are given in Table 2 together with the natural slope distribution. Table 2 illustrates that slope angle within ranges of 10–15° and 25–30° appear to be most conducive to landslide occurrence. However, we must emphasize that the reported locations of most of these landslides are normally the locations where the sliding earth materials stopped, not the initiation points or scars of landslide. Thus, we may interpret that most landslides stop at slopping ground with angles of 10–15° and 25–30°, and probably initiate at a larger slope angle.

Geology distribution of Hong Kong Island and landslides

In terms of the geology, Hong Kong has been subjected various geological actions and changes in geological history, therefore it is difficult to classify surface geology accurately. All igneous, volcanic, sedimentary and metaphoric rocks have been found in Hong Kong. There are probably more than 100 types of surface geology, depending on how we classify them. For the case of Hong Kong Island, the situation is little bit simpler. In this study, only five geological category have been adopted, and they are volcanics (mainly tuff), minor intrusives (including aplite, trachit, fledsparphyric rhyolite), sedimentary rocks and metasediments (sandstones, conglomerate, siltstone, mudstone and chert), granitic rocks (including granite and granodiorite), and debris deposits (talus, alluvium and colluvium). They are denoted in this study as Categories I, II, III, IV and V respectively. In terms of geological category, the percentages of landslides occurred within the Categories I–V are shown in Table 3 together with the natural geological distributions. The odds of landslide occurrence in Category IV (or for granitic rocks) is higher than 1 (odds equals 1 meaning a 50–50 chance of having landslide). Table 3 demonstrates that granitic formation is more conducive to landslide occurrence on Hong Kong Island.

Table 3 The distribution of geological category of Hong Kong Island and the distribution of 1,834 landslide events among five different geological categories

Geological category	Lithologic distribution (%)	Landslide distribution (%)
I	48	46
II	4	1
III	7	4
IV	27	38
V	14	11

I: volcanics; II: minor intrusives; III: sedimentary rocks and metasediments; IV: granitic rocks; V: debris deposits

Logistic Regression Analysis for the landslide data

In this section, we use logistic regression to correlate the landslide data with the geology, slope angle, slope aspect, and elevation using the standard statistical software (SPSS). A brief introduction of logistic regression analysis will first be given, following by the logistic regression analysis for the whole Hong Kong, and finally separate logistic analyses for northern and southern Hong Kong. The susceptibility maps by different approaches will be compared, and regional bias of the landslide data will be discussed.

Brief introduction of logistic regression

Various methods have been proposed to generate the landslide susceptibility maps, including neural network (Lee et al. 2004c), logistic regression (Rowbotham and Dudycha 1998; Dai and Lee 2002; Lee and Min 2001; Chau et al. 2004b; Ohlmacher and Davis 2003), wetness index approach (Gritzner et al. 2001), and numerical simulation (Chau and Lo 2004), and remote sensing (Lee et al. 2004b; Sarkar and Kanungo 2004). However, the most mature and well-adopted approach appears to be logistic regression analysis in relating landslide probability to local geology and geomorphology. Since our focus of this paper is not to investigate which mathematical or statistical method fit the landslide occurrence best, logistic regression would be used as a tool. We, however, expect that the finding on the regional bias of landslide data should not be too sensitive on the particular method used (in this case the logistic regression).

Logistic regression relates the term "odds" or "odds ratio", which is defined as the probability of landslide divided by the probability of no landslide (odds=Pr/(1−Pr), where Pr is the probability of landslide occurrence having values from 0 to 1). Therefore, if the odds is from 0 to 1 (corresponding to Pr from 0 to 0.5), the chance of landslide is less than that of non-occurrence (note that Pr=0.5 implies equal chance of landslide and non-landslide). For higher chance of landslide occurrence, Pr is from 0.5 to 1 or odds from 1 to ∞. Therefore, the odds is not a symmetric function with respect to the borderline of 1 because $0 < \text{odds} \leq 1$ for higher odds of non-occurrence whereas $1 < \text{odds} < \infty$ for higher odds of occurrence. One way to achieve symmetric property is to take the natural logarithm of odds and the resulting product is called "logit" or Z (i.e. Z=log(odds) or odds = exp(Z)), with $-\infty<Z<0$ for higher odds of non-occurrence and $0<Z<\infty$ for higher odds of occurrence. It is straightforward to invert Pr in terms of odds and then substitute the definition of Z into the resulting equation to yield

$$\text{Pr} = \frac{e^Z}{1 + e^Z} = \frac{1}{1 + e^{-Z}} \tag{1}$$

where Z is normally assumed to depend on the independent variables affecting landslide occurrence. In logistic regression analysis, the logit Z is assumed as a linear combination of the independent variables X_i (i=1, 2,. . ., n) as

$$Z = B_0 + B_1 X_1 + B_2 X_2 + \cdots + B_n X_n \tag{2}$$

where B_i are the coefficients reflecting the contribution of independent variables X_i to the logit Z. This particular form of dependence is also called log-linear. In logistic regression, the B_i are estimated using the maximum likelihood methods (in contrast to the least squares methods in linear regression). That is, the coefficients B_i that make our observed results most "likely" are selected. Normally, an iterative algorithm is used for B_i estimations. In the present analysis, the standard software SPSS is used to estimate B_i.

Note that the probability of landslide occurrence is a nonlinear function of each independent variable X_i because of the exponential function involved. The relative importance of different independent variables X_i can be examined easily. For example, from the following equation

$$e^Z = e^{B_0} e^{B_1 X_1} e^{B_2 X_2} \cdots e^{B_n X_n} \tag{3}$$

it is clear that odds increases exponentially with the independent variable X_i. Mathematically, if an independent variable X_i increases one unit, the odds of the occurrence increases by exp(B_i) times. Thus, a positive and large value of B_i indicates that a strong dependence of the odds in terms of the independent variable X_i. Logistic regression is good for analysis of categorical variables, that is, the variables can be divided into categories. For example, if slope angle is a variable affecting landslide occurrence, it is assumed dividable into several categories, such as 0–10°, 10–15°, 15–20°, 20–25°, 25–30°, 30–35°, 35–40° and ≥40°. Comparing to ordinary least square regression (or linear regression), logistic regression does not assume linearity of relationship between the independent variables and the dependent, does not require normally distributed variables, does not assume variables having equal statistical variances, and in general has less stringent requirements. Normally distributed error terms are not assumed. But as any other statistical methods, there are drawbacks. All relevant variables in the regression model must be included, but on the other hand all irrelevant variables must be excluded. Logistic regression is particularly useful for fitting data with discrete outcome, like in the present case that landslides either occur or do not occur (1 or 0 is normally assigned to them respectively).

For more details on the theory and concept of logistic regression, the readers are referred to Pampel (2000), Kleinbaum and Klein (2002), Hosmer and Lemeshow (2000), and Menard (2002).

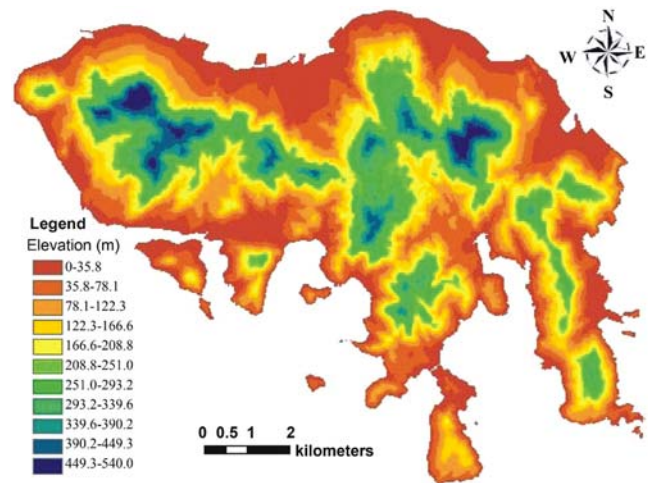### Logistic regression for the whole Hong Kong Island

In this study, the four chosen independent variables are geological condition, slope angle, slope aspect and elevation. As shown in Table 4, the geological is divided into 5 categories (i.e. I, II, III, IV and V for volcanics, minor intrusives, sedimentary rocks and metasediments, granitic and superficial deposits); slope angle into 8 categories (0–10°, 10–15°, 15–20°, 20–25°, 25–30°, 30–35°, 35–40° and ≥40°); slope aspect into 9 categories (North, Northwest, Northeast, South, Southwest, Southeast, East, West, Flat); and elevation into 11 categories (0–50 m, 50–100 m, 100–150 m, 150–200 m, 200–250 m, 250–300 m, 300–350 m, 350–400 m, 400–450 m, 450–500 m, and >450 m). Therefore, there are a total of 33 categories (i.e. n=33 in Eqs. (1) and (2)). As mentioned earlier, SPSS is used to get the B coefficients. Note that one category must be left out in the SPSS analysis. More specifically, we

Table 4 The variables, categories and coefficients of the logistic regression using all landslide data, landslide data of northern Hong Kong Island and landslide data of southern Hong Kong Island
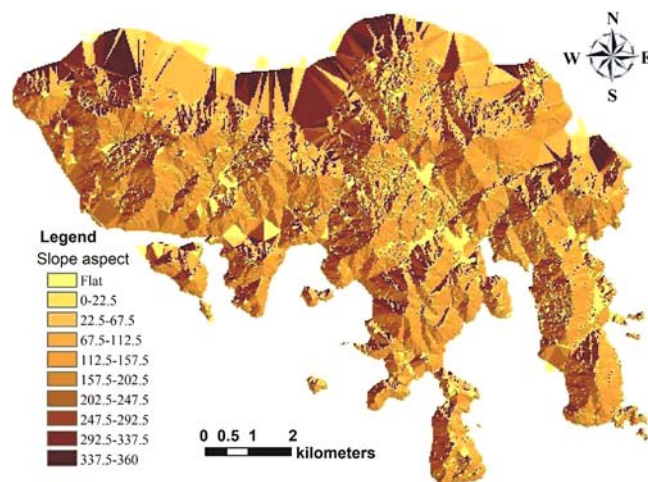
Variable	Category	B coefficients		
		All data	Northern data	Southern data
Elevation	0–50 m	2.521	2.7	2.309
	50–100 m	1.642	1.748	1.483
	100–150 m	0.749	1.144	−0.034
	150–200 m	0.061	−0.369	−0.661
	200–250 m	−0.732	−0.339	−1.765
	250–300 m	−0.877	−1.549	−0.128
	300–350 m	−1.01	−0.819	−1.183
	350–400 m	0.09	−0.078	−0.302
	400–450 m	0.599	0.064	1.576
	450–500 m	−1.427	−1.746	−0.719
	>500 m	−1.616	−1.494	−0.576
Slope aspect	Flat	−1.011	−1.004	−1.717
	North	0.207	−0.027	0.446
	Northeast	−0.225	−0.327	−0.055
	East	−0.096	−0.02	−0.288
	Southeast	−0.121	0.115	0.042
	South	0.327	0.283	0.156
	Southwest	0.572	0.63	0.717
	West	0.217	0.353	0.26
	Northwest	0.13	−0.003	0.439
Slope angle	0–10°	−0.088	−0.294	0.416
	10–15°	0.198	−0.04	0.91
	15–20°	0.346	0.128	0.631
	20–25°	−0.047	0.169	−0.234
	25–30°	−0.104	−0.05	−0.382
	30–35°	−0.16	−0.094	−0.431
	35–40°	−0.027	−0.001	−0.205
	>40°	−0.118	0.182	−0.705
Geology	I	0.32	0.457	0.237
	II	−0.778	−0.821	−0.597
	III	0.32	0.137	1.181
	IV	0.41	0.427	−0.009
	V	0.272	−0.2	−0.812
Constant	B_0	−1.008	−1	−0.615

Fig. 3 The raster plot of the elevation map generated from TIN model by using ArcGIS with a resolution of 30 m × 30 m. The lower elevation are denoted by orange and yellow, whereas higher grounds by green and blue

we can use "3D-Analyst" to convert the "TIN to Raster". The raster maps can be generated for elevation, slope aspect and slope angle, as shown in Figs. 3–5. The raster map of geology is generated from the 1:20,000 digital geology map of Hong Kong, as shown in Fig. 6. The resolution of the DEM model and other raster maps is 30×30 m, which is found to be a good choice by Lee et al. (2004a). The digital version of the geological map is supplied by GEO. The raster value of slope angle, geology category, slope aspect, and elevation are extracted for each landslide point using "zonal statistics" function in ArcGIS. These geological and geomorphological data are not only generated for the 1,834 landslides but also for another 1,834 non-landslide points (this is a requirement by SPSS), which are selected randomly throughout the region (excluding flat lands) but spatial overlapping with the landslide points must be avoided (see Fig. 7). In reality, the generation of these non-landslide points can also be done using ArcCatalog and ArcMap. Once this non-landslide shape file is available, we can use "zonal statistics" to extract the required information of geology, slope

can either use the binary coding or dummy variable coding (i.e. set the value of the right category of a variable as 1 and all other zeros with the last category left out as the reference category for the variable), or the effect or deviation coding (i.e. set the value of the last category to −1). For the case of deviation coding, the B coefficient of the reference or last category can be obtained by taking the negative value of the sum of all other B coefficients in the same variable; and for the case of dummy variable coding, the B coefficient of the reference category is zero. Although the resulting B values are not the same, the final susceptibility maps generated from these coefficients are the same. In the present approach, the effect coding is used since the logistic regression coefficient of the last category can also be displayed.

To generate the geological and geomorphological data for logistic regression analysis, ArcGIS has been used. In particular, the 1:5,000 digital contour map of 10 m elevation interval is used to generate the TIN (Triangulated Irregular Network) and DEM (Digital Elevation Model), which is shown in Fig. 3. On the basis of this TIN file, raster (or pixel in the digital map) files of elevation, slope aspect and slope angle can be generated for all locations of Hong Kong Island. In particular,

Fig. 4 The raster plot of slope aspects generated from TIN model by using ArcGIS with a resolution of 30 m × 30 m. The aspect directions facing northeast are lighter in color whereas those facing northwest are darker and all values are between 0 and 360°
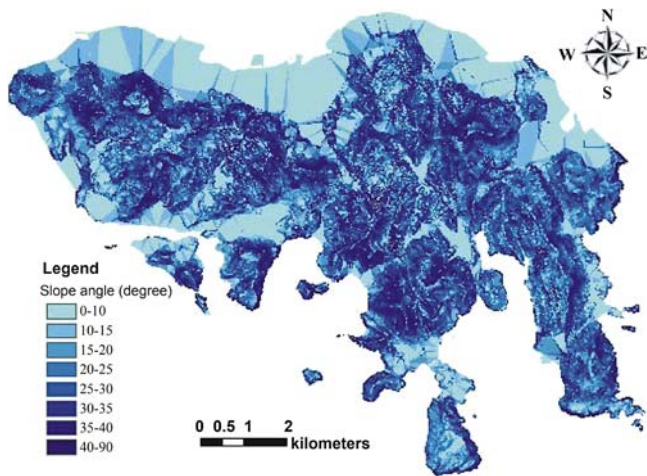
Fig. 5 The raster plot of slope angle generated from TIN model by using ArcGIS with a resolution of 30 m × 30 m. Steeper slopes are shown in dark blue whereas more gentle slopes by light blue, and all values are from 0 to 90°. Note that the artificial triangular sections on the north coastline are results from the TIN generated process from the contour map
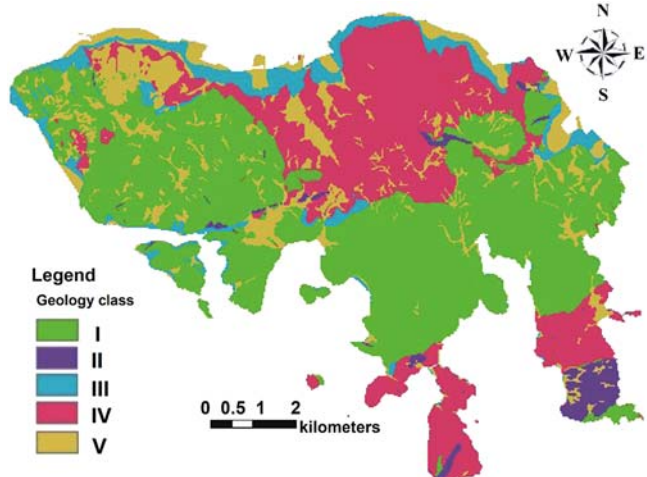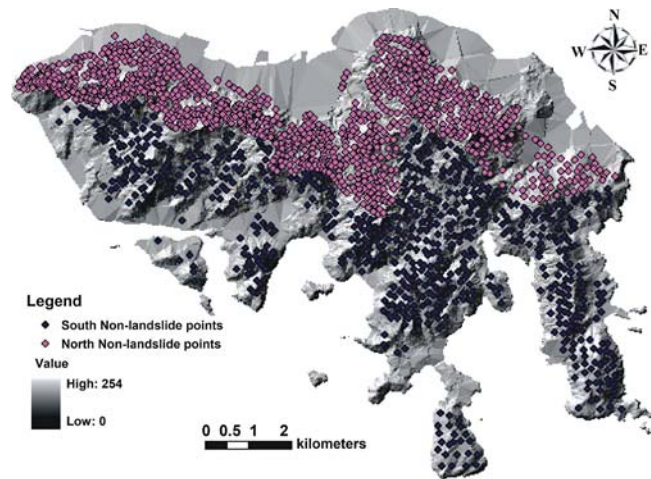

Fig. 7 A Hong Kong Island map showing non-landslide incident points used in our logistic regression analysis (a total of 1,834 points). The chosen non-landslide points are pink and blue for those lying north and south of the divide line respectively. Note that we deliberately do not choose points from the reclamation areas that the landslide susceptibility is minimal

Fig. 6 The raster plot of geological category generated from TIN model by using ArcGIS with a resolution of 30 m × 30 m. Categories I, II, III, IV and V are for volcanics, minor intrusives, sedimentary rocks and metasediments, granitic rocks and debris deposits

angle, slope aspect and elevation, similar to what we have done for the landslide points. The extracted data from ArcMap can be read and edited by using Microsoft Excel, before converted to database format for SPSS analysis.

For SPSS analysis, a data file containing all landslide data and non-landslide data and the associated values of variables (i.e. put 1 into the "right" category of each slope aspect, slope angle, elevation and geology) must be prepared in a worksheet format and all information in this file will be used simultaneously in the SPSS analysis. Note that nearly all data in this file can be extracted from the "Zonal Statistics" function under "Spatial Analyst" in ArcMap.

Once the file is prepared, SPSS can be used to estimate the *B* coefficient. The result of the analysis is shown in column 3 of Table 4. It can be seen that elevation of 0–50 m is the most dominant factor in the landslide occurrence. As mentioned earlier, all *B* coefficients larger than 0 implies a probability higher than 0.5 whereas

negative *B* coefficients suggest a landslide occurrence of less than 50% chance. Therefore, in the statistical sense granitic formation at an elevation between 0 and 50 m with a slope angle of 15–20° and slope aspect facing southwest is more susceptible to landslide occurrence.

Logistic regression for the northern Hong Kong Island

As discussed in the Introduction section, the main objective of this study is to examine the regional bias of landslide data used in predicting the landslide susceptibility. Therefore, this section examines the possibility of the regional bias of the landslide data. The Hong Kong Island is divided into two parts using a roughly west-east trending water divide shown in Fig. 1 such that the northern part mainly consists of north-facing slopes whereas the southern part mainly consists of south-facing slopes. The total number of landslide events in the northern Hong Kong Island is of 1,115 landslides. Another set of 1,115 non-landslide points are selected and geological and geomorphological information of these data points are also extracted by "zonal statistics" function of ArcMap. The locations of these points are selected such that these points are distributed uniformly over the northern Hong Kong Island, which are shown in Fig. 7.

The results of the landslide data analysis of northern Hong Kong Island are shown in column 4 of Table 4. It was found that the elevation influence on the landslide occurrence is stronger than that for all landslide data (results in column 3 of Table 4). Most landslides occur on slope of steeper than 40°. The type of geology most conducive to landslides is volcanic or tuff.

Logistic regression for the southern Hong Kong Island

The slopes in the southern part of the west-east trending water divide are, in general, more gentle than the slopes of the northern parts, and the area is less developed. The southern Hong Kong Island is facing the South China Sea to the south and the Pacific Ocean to the east. The total number of landslide events in the southern Hong Kong Island is 719, and again the same number non-landslide points are selected uniformly over the area (see Fig. 7). The results of the analysis are given in column 5 of Table 4. As shown in Table 4, elevation of 0–100 m remains the most dominant factor in the landslide occurrence. But comparing to the northern Hong Kong Island, landslide is more

likely to occur at elevation of 400–450 m with a *B* coefficient of 1.576 comparing to 0.064 for the northern part. A closer look of the landslide data reveals that many landslides occur in the "developed peak area" occurred south of the water divide, resulting in a higher value of *B* coefficient. In terms of the slope angles, the southern Hong Kong Island landslides are more likely to occur for more gentle slopes whilst for the northern Hong Kong Island landslides are also likely to occur at slope angles larger than 40° (see Table 4). As shown in Fig. 1, the slopes of the northern Hong Kong Island are steeper than the southern part; therefore, as expected, the probability of landslide occurrence on steeper slopes is higher. In terms of geology, both tuff and granite are conducive to landslides in the north whereas tuff and sedimentary rocks appear to be more conducive to landslide occurrence in the south.

By comparing the results of logistic regression analyses for both south and north Hong Kong Island, it appears that there is a strong regional dependence of the landslide data (see the difference in the *B* coefficient given in Table 4). More discussion on this will be given again in later sections.

Combined hazard map and its reliability

To illustrate the difference between the logistic regression analyses in the sections Logistic regression for the whole Hong Kong Island, logistic regression for the northern Hong Kong Island and logistic regression for the southern Hong Kong Island, the probability of landslide can be estimated by substituting the *B* coefficients given in Table 4 into Eqs. (1) and (2). For the present case, *n* equals 33 for X_i (*i*= 1, ..., *n*) (see Table 4) and with *i*=1, ..., 11 for elevations, *i*=12, ..., 20 for slope aspects, *i*=21,..., 28 for slope angles, and *i*=29,..., 33 for geology. To illustrate how to find the probability of landslide occurrence, we consider a particular example. For example, for a location with elevation of 35 m, slope aspect is facing north, slope angle of 23°, and geology of granitic formation, we can set X_2, X_{13}, X_{24} and X_{32} equals 1 and all others X_i set to zero in Eq. (2), then the probability can be estimated from (1).

The actual calculation is done by using the "Raster Calculator" function under the "Spatial Analyst" of ArcMap. The final map is normally generated using "Raster calculator" a few times. The raster calculation is normally done on one raster layer (or one variable) by one raster layer. We first assign the *B* values to each of the category of each variable. After B value assignment to each category, we can combine all layers by using Eqs. (1) and (2) to yield the susceptibility using raster calculator. The final logit is then calculated before using Eq. (1) to find the susceptibility. In addition, since the reclaimed flat land on the northern edge of the Hong Kong Island is far from hillside, the landslide hazard in this flat areas have to be excluded. In the present analyses, the hazard will be set to zero if any location has an elevation of smaller than 20 m and a slope angle of less than 5°. In all hazard maps given in Figs. 8–11, all unrealistic landslide hazards on flat ground far from hillside has been removed using such approach.

The hazard maps estimated by using the *B* coefficients from the northern data and from the southern data are given in Figs. 8 and 9 respectively. Both landslide susceptibility maps can be compared with the original landslides data shown in Fig. 2. It can be seen that if the formula for predicting the northern region is extended to estimate the whole Hong Kong island, the prediction will differ significantly from the actual landslide data in the extended part. The same is also true for extending the southern formula to the northern area.

Figure 10 shows the landslide susceptibility map by combining the northern map with the southern map (i.e. combining the northern
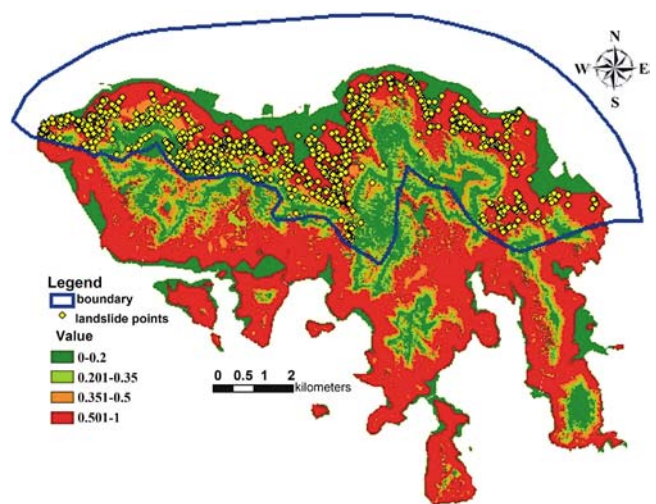

Fig. 8 Landslide susceptibility map of the whole Hong Kong Island generated using the *B* coefficients from landslide data in the Northern part of HK Island only (see the fourth column of Table 4). The original landslide locations are denoted as solid yellow rhombus. The red zone represents landslide susceptibility of higher than 0.5. All given values are landslide probability predicted by Eqs. (1) and (2) using *B* values in column 4 of Table 4
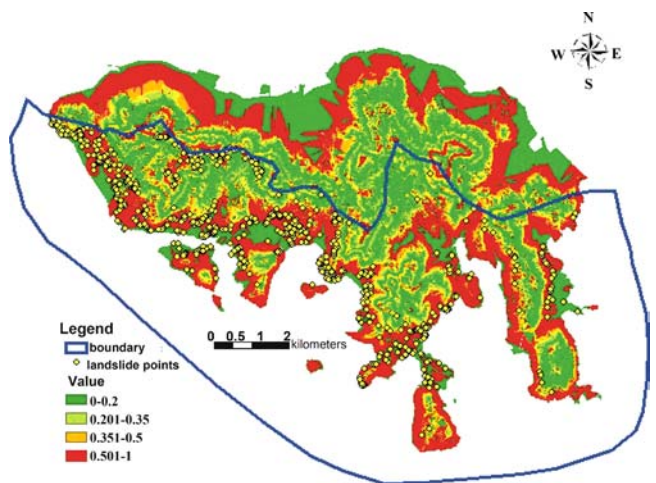

Fig. 9 Landslide susceptibility map of the whole Hong Kong Island generated using the *B* coefficients from landslide data in the Southern part of HK Island only (see the fifth column of Table 4). The original landslide locations are denoted as solid yellow rhombus. The red zone represents landslide susceptibility of higher than 0.5. Note that the red zones for the "Southern Hong Kong" are substantially smaller than that predicted in Fig. 8. All given values are landslide probability predicted by Eqs. (1) and (2) using *B* values in column 5 of Table 4

part of Fig. 8 and the southern part of Fig. 9). As it can be shown in Fig. 10 that the landslide points are mainly within the red zone with a probability of landslide occurrence of larger than 50%. Therefore, if we set 0.5 as the probability of the borderline between a high and low hazard zone, the susceptibility map of Fig. 10 appears to be acceptable and can reflect what had occurred in the past.

A standard analysis in logistic regression normally includes a so-called "classification plot of 0 and 1" for the observed groups and the predicted probability. In particular, the probability of landslide occurrences can be calculated for all raster points of landslide occurrence and of non-landslide occurrence. If the predicted probability of any landslide raster point is larger than 0.5 (a number of 1 is
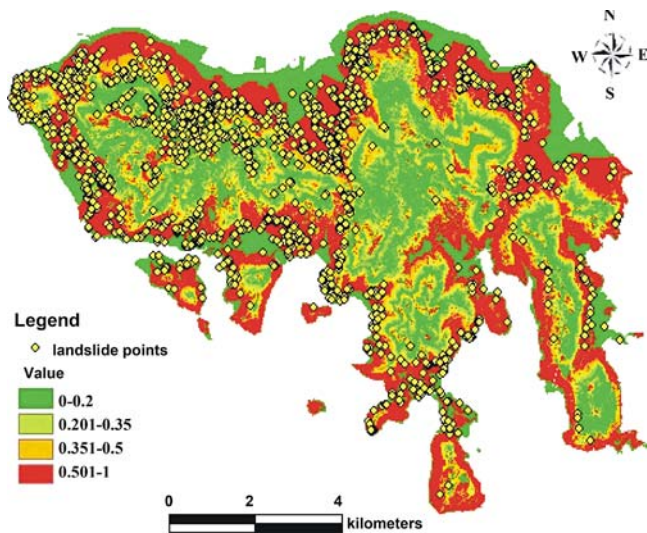
Fig. 10 The landslide susceptibility map of the whole Hong Kong Island by combing the northern susceptibility map from the "north-data-analysis" and the southern susceptibility map from the "south-data-analysis". The map is a result of combining the northern part of Fig. 8 and the southern part of Fig. 9. The original landslide points are given as yellow dots
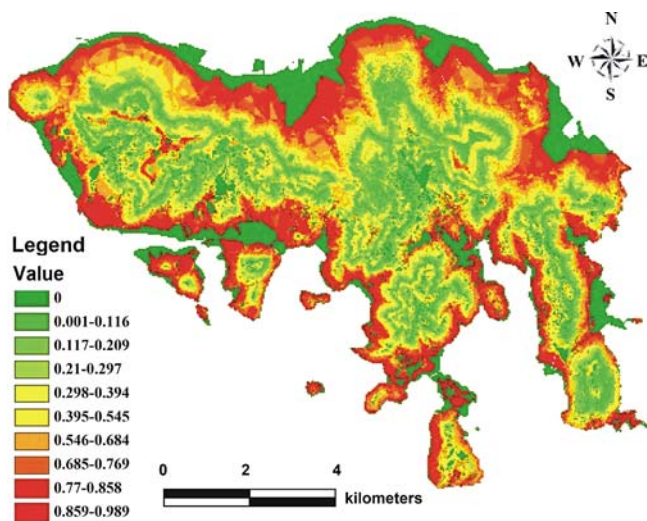

Fig. 11 A more refined landslide susceptibility map of the whole Hong Kong Island shown in Fig. 10, using nine equal increments of susceptible levels. The red zones are more susceptible to landslide occurrence

Table 5 Comparison of predicted areas with susceptibility larger than 0.5 using all-data-analysis (using all landslide data), south-data-analysis (using landslide data south of the water divide in Fig. 2) and north-data-analysis (using landslide data north of water divide in Fig. 2)

Region	Area with susceptibility ≥0.5 (km²)		
	Using north data	Using south data	Using all data
North	13.93	9.76	12.68
South	24.17	16.01	19.90

≥0.5 (or "red zone") in Figs. 8–10, the difference between different analyses is much bigger. Table 5 compiles the areas of the "red zones" of both the north and south regions from different logistic regression analyses using south, north and all data (given in Figs. 8, 9 and 10 respectively). There is a big difference in terms of the size of the predicted susceptible zone. For example, the size of the susceptibility zone (with probability ≥0.5) of south Hong Kong obtained by all-data-analysis and north-data-analysis were 24% and 51% larger than that of the south-data-analysis. This strongly indicates that there is a strong regional bias. Therefore, the standard classification plot does not appear to be an effective way of investigating the regional bias of data.

We also have conducted another series of logistic regression analysis by removing each of the categories of elevation, slope aspect, slope angle, and geology. This is a standard technique in logistic regression analysis to examine the statistical significance of each of the selected independent variables (e.g. Hosmer and Lemeshow 2000; Menard 2002). The "−2 log likelihood ratio" (or the so-called −2LL) for each of the all, northern and southern data analyses were obtained, and it was found that −2LL is the smallest if all variables of elevation, slope aspect, slope angle and geology are included. The −2LL can be considered as a goodness of fit of the model. Since the likelihood is between 0 and 1, log likelihood would be from negative infinity to zero, and in turn the −2LL would be from 0 to infinity. Therefore, the original analysis of incorporating all variables should be used.

In comparison to other studies, similar logistic regression by Dai and Lee (2002) for Lantau Island in Hong Kong also found that elevation is the most dominant factor, among slope aspect, lithology, land-use type and slope angle. Since the population of Lantau Island is much less than that of Hong Kong Island, therefore it seems that the elevation dependence is not caused by the density of human activities (say the existence of cut slopes behind roads and buildings) alone, but appears to be an inherent characteristics of landslide of Hong Kong. However, this observation needs to be further verified. One way to verify this is to add cut slope angle as a new independent variable and repeat the analysis to see whether the elevation dependence of both Lantau Island and Hong Kong would disappear. However, such analysis is out of the scope of this study.

Another way to investigate the reliability of the hazard analysis is to consider the ratio of the landslide location points versus the landslide susceptibility (Lee and Min 2001). According to Lee and Min (2001), the susceptibility or the landslide probability should first be divided into intervals such that within each of them the amount of coverage area is about the same. Such a plot of the susceptibility map is given in Fig. 11. The results are also summarized in Table 6. The area with susceptibility =0 is resulted from our exclusion analysis for reclamation areas (see earlier section). The rest of the areas are roughly divided into zone of 7 km² (see columns 2 and 3 of Table 6), according to the level of susceptibility. The observed landslide events within each of this zone are extracted by using "Zonal Statistics" of ArcMap and are reported in columns 4 and 5. Note that $b/a>1$

assigned to this point), we consider the prediction is correct, and the same for non-landslide raster point with a predicted probability of less than 0.5 (a number of 0 is assigned to this point). Then, all data points (both landslide and non-landslide points) can be used as a check for the coherence of the hazard map. In particular, we can count the number of "1" in the zone with probability larger than 0.5 and the number of "0" in the zone with probability less than 0.5. It was found that the correct percentages are 74.4%, 73.7% and 79.3% for the cases of using all data points (1,834×2=3,668), northern data points (1,115×2=2,230), southern data points (719×2=1,438) respectively. The correct percentage of the combined hazard map is 76.5 which is only slightly higher than that for using all data. However, if we measure the size of the zone with susceptibility

Table 6 Comparison of landslide occurrence and landslide susceptibility map using logistic regression method for Hong Kong Island

Range of susceptibility	Area distribution		Landslides occurred		b/a	Landslide (km^2)
	Area (km^2)	Area ratio (a, %)	Landslide number	Area ratio (b, %)		
0.00	14.7141	18.570	182	9.924	0.534	12.369
0.001–0.116	7.0317	8.874	23	1.254	0.141	3.271
0.117–0.209	7.1379	9.008	55	2.999	0.333	7.705
0.21–0.297	7.2801	9.188	93	5.071	0.552	12.774
0.298–0.394	7.5195	9.490	102	5.562	0.586	13.565
0.395–0.545	7.2072	9.096	168	9.160	1.007	23.310
0.546–0.684	7.2936	9.205	262	14.286	1.552	35.920
0.685–0.769	6.8886	8.694	338	18.430	2.120	49.067
0.77–0.858	6.9786	8.807	278	15.158	1.721	39.836
0.859–0.989	7.1856	9.069	333	18.157	2.002	46.343

The total area of Hong Kong Island is 79.2369 km^2. The total number of landslides is 1,834. When $b/a=1$, the susceptibility of landslide is average. The average landslide for Hong Kong Island is 23.14 landslides/km^2. The area ratio, a, represents the distribution of area within each susceptibility given in column 1. The area ratio, b, represents the distribution of area containing at least one landslide for each of the susceptibility range in column 1 (counting using cells of 30 m × 30 m)

indicates a chance of landslide occurrence higher than average. As shown in Table 6, the susceptibility and probability of larger than 0.5 roughly corresponds to $b/a>1$. This further indicates that the present logistic regression does yield a reliable hazard or susceptibility map for Hong Kong Island.

Another way to interpret the same data is to consider the landslide events per square kilometer, comparing to the average number of landslides per square kilometer (see the last column of Table 6). Note that the first row of Table 6 is 0 for the susceptibility, which was set artificially to zero for flat reclamation areas (as remarked earlier). Therefore, the first row of data is somewhat artificial. If we compare the second and the last row of the susceptibility (i.e. consider susceptibility of 0.001–0.116 and 0.859–0.989), the landslide per square kilometer increases from 3.271 to 46.343, or an increase of 14 times. This value is about the same as 15.8, which is the ratio of the median of the second and the last row of 15.8 (i.e. 0.924/0.0585). Therefore, the predicted susceptibility agrees very well with the actual landslide observation.

To further examine this comparison, Table 7 compiles the median susceptibility level of each range (column 1 in Table 6) versus the actual observed landslide density within each range. To make comparison, the third and fifth columns give the normalized median susceptibility versus the normalized observed landslide density. The last column of Table 7 shows that the difference between the predictions (column 3) and observations (column 5) is less than 30% for all susceptibility levels. Therefore, the present logistic regression model is considered quite reliable.

Discussions and conclusion

A framework for analyzing landslide hazard analysis using GIS is proposed for Hong Kong Island on the basis of 1,834 landslide records from 1982 to 2002. Logistic regression analysis was conducted to yield a susceptibility map for Hong Kong Island. In addition, to examine the possible regional dependence of the landslide data, the Hong Kong Island is divided into the southern and the northern regions by an east-west trending water divide. Although the north region is smaller than the southern, 1,115 landslide falls onto the northern part

Table 7 Comparison of median landslide susceptibility with the observed landslide density at various susceptibility levels

Range of susceptibility	Median susceptibility		Observed landslide density (number/km^2)		% difference[a]
	Actual	Normalized	Actual	Normalized	
0.001–0.116	0.0585	1	3.271	1	–
0.117–0.209	0.163	2.78	7.705	2.34	−16
0.21–0.297	0.2535	4.33	12.774	3.91	−10
0.298–0.394	0.346	5.91	13.565	4.15	−30
0.395–0.545	0.47	8.03	23.310	7.13	−11
0.546–0.684	0.615	10.51	35.920	10.98	4
0.685–0.769	0.727	12.43	49.067	15.00	21
0.77–0.858	0.814	13.91	39.836	12.18	−12
0.859–0.989	0.924	15.79	46.343	14.17	−7

[a] This is the % difference between columns 3 and 5

and the remaining 719 landslide data falls within the southern region. That is, the northern Hong Kong has a higher density of landslide occurrence. Different susceptibility analyses have been conducted by using the north data, the south data and the whole set of data separately in generating the susceptibility maps. It was found that the susceptibility map of the southern Hong Kong Island generated by using north data differs significantly from that generated by using south data only, and similarly for the northern Hong Kong Island. Therefore, a final version of susceptibility map is generated by combining the northern susceptibility map with the southern one; and this final map appears to reflect closely the spatial distributions of historical events. The elevation appears to be the most dominant factor in controlling landslide occurrence. More refined comparisons at all levels of susceptibility are given in Table 7. The goodness of fit of the model was also examined by using the relative values of −2LL. Classification plot and occurrence ratio proposed by Lee and Min (2001) were used to examine the reliability of the proposed hazard map as well. The results of these goodness-of-fit analyses suggest that the present hazard model for Hong Kong Island is reliable and should at least provide a good first order prediction of landslide occurrence. But, of course, more sophisticated approach by incorporating landslide dynamics should be used if a more accurate prediction is needed (e.g. Chau and Lo 2004).

In addition, note that the majority of the landslide inventory used in this study is for cut slope failures. If we include a new variable "cut slope", add the cut slope angle as category and repeat the logistic regression analysis, the landslide susceptibility for low areas without cut slopes probably would be much lower. The strong correlation of landslide susceptibility with low elevation may subsequently decrease. However, the inventory of the 57,000 cut slopes in Hong Kong and their corresponding cut slope angles are not readily available in GIS format. This revised analysis should be a meaningful analysis in the future.

An obvious drawback in our present logistic regression model is that the landslide susceptibility of flat areas has to be removed manually whereas landslide susceptibility maps by other methods, such as Discriminant Analysis (Santacana et al. 2003), Artificial Neural Networks (Lu and Rosenbaum 2003), Matrix Method (Fernandez et al. 1999), and Favourability Functions (Remondo et al. 2003), appear not having this problem.

One practical and theoretical question that was frequently asked is how much data is needed for developing a reasonable logistic regression model. In the statistical sense, more landslide data should

yield more reliable model (if there is a strong correlation). Although it is difficult to give a definite minimum number of data needed for such logistic regression analysis, we believe that for the present case of Hong Kong Island (i.e. an average of 23 landslide/km²) the number of landslide events should provide a rough guideline for similar analysis for other parts of the world.

Nevertheless, this study strongly illustrates that the historical landslide data may be bias and affected by human activities and geological settings on a regional basis. Therefore, indiscriminate use of regional-biased data should be avoided. Only regional landslide records for similar level of human activities and similar geological setting should be analyzed simultaneously. More importantly, landslide statistics should not be borrowed indiscriminately from one region to another in generating landslide susceptibility maps.

References

Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65(1/2):15–31

Ayalew L, Yamagishi H, Ugawa N (2004) Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. Landslides 1(1):73–81

Cancelli A, Crosta G (1994) Hazard and risk assessment in rockfall prone areas. In: Skipp BO (ed) Risk and reliability in ground engineering. Thomas Telford, London, pp 177–190

Chau KT, Lo KH (2004) Hazard assessment of debris flow for Leung King Estate of Hong Kong by incorporating GIS with numerical simulations. Nat Hazards Earth Syst Sci 4(1):103–116

Chau KT, Wong RHC, Lui J, Lee CF (2003) Rockfall hazard analysis for Hong Kong based on rockfall inventory. Rock Mech Rock Eng 36(5):383–408

Chau KT, Sze YL, Fung MK, Wong WY, Fong EL, Chan LCP (2004a) Landslide hazard analysis for Hong Kong based on landslide inventory and GIS. Comput Geosci 30:429–443

Chau KT, Tang YF, Wong RHC (2004b) GIS based rockfall hazard map for Hong Kong. Int J Rock Mech Mining Sc 41(3):530

Coppock JT (1995) GIS and natural hazards: An overview from a GIS perspective. In: Carrara A, Guzzetti F (eds) Geographical information systems in assessing natural hazards. Kluwer, Dordrecht, pp 21–34

Corominas J, Santacana N (2003) Stability analysis of the Vallcebre translational slide, Eastern Pyrenees (Spain) by means of a GIS. Nat Hazards 30(3):473–485

Corominas J, Copons R, Vilaplana JM, Altimir J, Amigó J (2003) Integrated Landslide Susceptibility Analysis and Hazard Assessment in the Principality of Andorra. Nat Hazards 30(3):421–435

Dai FC, Lee CF (2001) Terrain-based mapping of landslide susceptibility using a geographical information system: A case study. Can Geotech J 38(5):911–923

Dai FC, Lee CF (2002) Landslide characteristics and, slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42(3/4):213–228

Dai FC, Lee CF (2003) A spatiotemporal probabilistic modelling of storm-induced shallow landsliding using aerial photographs and logistic regression. Earth Surf Process Landforms 28(5):527–545

Dai FC, Lee CF, Li J, Xu ZW (2001) Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. Environ Geol 40(3):381–391

Dai FC, Lee CF, Ngai YY (2002) Landslide risk assessment and management: An overview. Eng Geol 64(1):65–87

Einstein HH (1988) Landslide risk assessment procedure. In: Proceedings of the 5th international symposium on landslides, vol 2, Balkema, Rotterdam, pp 1075–1090

Einstein HH (1997) Landslide risk—Systematic approaches to assessment. In: Cruden D, Fell R (eds) Landslide risk assessment. Balkema, Rotterdam, pp 25–50

Fell R, Hartford D (1997) Landslide risk management. In: Cruden D, Fell R (eds) Landslide risk assessment. Balkema, Rotterdam, pp 51–109

Fernandez CI, Del Castillo TF, El Hamdouni R, Montero JC (1999) Verification of landslide susceptibility mapping: A case study. Earth Surf Process Landforms 24(6):537–544

Fyfe JA, Shaw R, Campbell SDG, Lai KW, Kirk PA (2000) The quaternary geology of Hong Kong. Geotechnical Engineering Office, Hong Kong, 210 pp

Geotechnical Engineering Office (GEO) (1996) Compilation of a database on landslide consequence. Submitted by Mitchell, McFarlane, Brentnall & Partners Int. Ltd., Agreement No. GEO 8/95, Final report for Special Project Division, GEO, 310 pp

Gritzner ML, Marcus WA, Aspinall R, Custer SG (2001) Assessing landslide potential using GIS, soil wetness modeling and topographic attributes, Payette River, Idaho. Geomorphology 37(1/2):149–165

Hansen A (1984) Landslide hazard analysis. In: Brunsden D, Prior DB (eds) Slope instability, Chapter 13. Wiley, New York, pp 523–602

Hosmer DW, Lemeshow S (2000) Applied logistic regression, 2nd edn. Wiley, New York, 373 pp

Hungr O (1997) Some methods of landslides hazard intensity mapping. In: Cruden D, Fell R (eds) Landslide risk assessment. Balkema, Rotterdam, pp 215–226

Kleinbaum DG, Klein M (2002) Logistic regression: A self-learning text. Springer-Verlag, New York, 513 pp

Larsen MC, Parks JE (1997) How wide is a road? The association of roads and mass-wasting in a forested mountain environment. Earth Surf Process Landforms 22:835–848

Lee CF, Ye H, Yeung MR, Shan X, Chen G (2001) AIGIS-based methodology for natural terrain landslide susceptibility mapping in Hong Kong. Episodes 24(3):150–159

Lee S (2004) Application of likelihood ratio and logistic regression models to landslide susceptibility mapping using GIS. Environ Manage 34(2):223–232

Lee S (2005a) Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data journals. Int J Remote Sens 26(7):1477–1491

Lee S (2005b) Application and cross-validation of spatial logistic multiple regression for landslide susceptibility analysis. Geosci J 9(1):63–71

Lee S, Min K (2001) Statistical analysis of landslide susceptibility at Yongin, Korea. Environ Geol 40(9):1095–1113

Lee S, Choi J, Woo I (2004a) The effect of spatial resolution on the accuracy of landslide susceptibility mapping: A case study in Boun, Korea. Geosci J 8(1):51–60

Lee S, Choi J, Min K (2004b) Probabilistic landslide hazard mapping using GIS and remote sensing data at Boun, Korea. Int J Remote Sens 25(11):2037–2052

Lee S, Ryu JH, Won JS, Park HJ (2004c) Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Eng Geol 71(3/4):289–302

Leroi E (1997) Landslide risk mapping: Problems, limitation and developments. In: Cruden D, Fell R (eds) Landslide risk assessment. Balkema, Rotterdam, pp 239–250

Lu P, Rosenbaum MS (2003) Artificial neural networks and Grey Systems for the prediction of slope stability. Nat Hazards 30(3):383–398

Luzi L, Pergalani F, Terlien MTJ (2000) Slope vulnerability to earthquakes at subregional scale, using probabilistic techniques and geographic information systems. Eng Geol 58(3/4):313–336

Marquinez J, Duarte RM, Farias P, Sanchez MJ (2003) Predictive GIS-based model of rockfall activity in Mountain Cliffs. Nat Hazards 30(3):341–360

Menard SW (2002) Applied logistic regression analysis, 2nd edn. Sage, Thousand Oaks, CA, 111 pp

Menendez-Duarte R, Marquinez J, Devoli G (2003) Slope instability in Nicaragua triggered by Hurricane Mitch: Distribution of shallow mass movements. Environ Geol 44(3):290–300

Miles SB, Keefer DK (1999) Evaluation of seismic slope-performance models using a regional case study. Environ Eng Geosci 6(1):25–39

Myster RW, Thomlinson JR, Larsen MC (1997) Predicting landslide vegetation in patches on landscape gradients in Puerto Rico. Landscape Ecol 12(5):299–307

Ohlmacher GC, Davis JC (2003) Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Eng Geol 69(3/4):331–343

Pampel FC (2000) Logistic regression: A primer. Sage, Thousand Oaks, CA, 86 pp

Refice A, Capolongo D (2002) Probabilistic modeling of uncertainties in earthquake-induced landslide hazard assessment. Comput Geosci 28(6):735–749

Remondo J, Gonzalez A, De Teran JRD, Cendrero A, Fabbri A, Chung CJF (2003) Validation of landslide susceptibility maps: Examples and applications from a case study in northern Spain. Nat Hazards 30(3):437–449

Rowbotham DN, Dudycha D (1998) GIS modelling of slope stability in Phewa Tal watershed, Nepal. Geomorphology 26(1–3):151–170

Saha AK, Gupta RP, Sarkar I, Arora MK, Csaplovics E (2005) An approach for GIS-based statistical landslide susceptibility zonation—with a case study in the Himalayas. Landslides 2(1):61–69

Santacana N, Baeza B, Corominas J, De Paz A, Marturia J (2003) A GIS-based multivariate statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet area (Eastern Pyrenees, Spain). Nat Hazards 30(3):281–295

Sarkar S, Kanungo DP (2004) An integrated approach for landslide susceptibility mapping using remote sensing and GIS. Photogrammetric Eng Remote Sens 70(5):617–625

Sassa K, Wang G, Fukuoka H, Wang F, Ochiai T, Sugiyama M, Sekiguchi T (2004) Landslide risk evaluation and hazard zoning for rapid and long-travel landslides in urban development areas. Landslides 1(3):221–235

Varnes DJ (1984) Landslide hazard zonation: A review of principles and practice. UNESCO, France, pp 1–63

Wang HB, Sassa K (2005) Comparative evaluation of landslide susceptibility in Minamata area, Japan. Environ Geol 47(7):956–966

Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Eng Geol 79(3/4):251–266

K. T. Chau (✉) · J. E. Chan
Department of Civil and Structural Engineering, The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong
e-mail: cektchau@polyu.edu.hk
Tel.: +852-27666015
Fax: +852-23346389