

Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS

Hossein Mojaddadi, Biswajeet Pradhan, Haleh Nampak, Noordin Ahmad & Abdul Halim bin Ghazali

To cite this article: Hossein Mojaddadi, Biswajeet Pradhan, Haleh Nampak, Noordin Ahmad & Abdul Halim bin Ghazali (2017) Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS, Geomatics, Natural Hazards and Risk, 8:2, 1080-1102, DOI: [10.1080/19475705.2017.1294113](https://doi.org/10.1080/19475705.2017.1294113)

To link to this article: <https://doi.org/10.1080/19475705.2017.1294113>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 01 Mar 2017.



Submit your article to this journal [↗](#)



Article views: 1109



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 5 View citing articles [↗](#)



Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS

Hossein Mojaddadi^a, Biswajeet Pradhan ^{a,b}, Haleh Nampak^a, Noordin Ahmad^a and Abdul Halim bin Ghazali^a

^aDepartment of Civil Engineering, Faculty of Engineering, University Putra Malaysia, Serdang, Selangor, Malaysia;

^bDepartment of Energy and Mineral Resources Engineering, Sejong University, Seoul, Republic of Korea

ABSTRACT

In this paper, an ensemble method, which demonstrated efficiency in GIS based flood modeling, was used to create flood probability indices for the Damansara River catchment in Malaysia. To estimate flood probability, the frequency ratio (FR) approach was combined with support vector machine (SVM) using a radial basis function kernel. Thirteen flood conditioning parameters, namely, altitude, aspect, slope, curvature, stream power index, topographic wetness index, sediment transport index, topographic roughness index, distance from river, geology, soil, surface runoff, and land use/cover (LULC), were selected. Each class of conditioning factor was weighted using the FR approach and entered as input for SVM modeling to optimize all the parameters. The flood hazard map was produced by combining the flood probability map with flood-triggering factors such as; averaged daily rainfall and flood inundation depth. Subsequently, the hydraulic 2D high-resolution sub-grid model (HRS) was applied to estimate the flood inundation depth. Furthermore, vulnerability weights were assigned to each element at risk based on their importance. Finally flood risk map was generated. The results of this research demonstrated that the proposed approach would be effective for flood risk management in the study area along the expressway and could be easily replicated in other areas.

ARTICLE HISTORY

Received 2 November 2016
Accepted 8 February 2017

KEYWORDS

Flood risk assessment; vulnerability; hazard; GIS; SVM; 2D HRS; inundation; remote sensing; Malaysia

1. Introduction

Flood events are typically regarded to be the most common natural disaster worldwide (Stefanidis & Stathis 2013). Hence, flood risk management is an important challenge in many cities. Rapid urbanization, population growth, economic development, and climate change will increase the magnitude of this challenge (Huong & Pathirana 2013). Considerable and irreparable damages to farmlands, transportation, bridges, and many other aspects of urban infrastructure prove the urgent requirement for flood control and prevention (Tehrany et al. 2014; Pradhan et al. 2014).

Flood risk analysis and risk mitigation are two components of flood risk management. Flood risk analysis aims to investigate where the risk of flood occurrence is unacceptably high and where risk mitigation actions are required. Therefore, comprehensive flood risk analysis by detecting hazardous and risky areas is an essential part of risk management to estimate the amount of damages that can occur because of flooding (Meyer et al., 2008).

The recent improvements in the efficiency of remote sensing (RS) and geographic information system (GIS) technologies have initiated a revolution in hydrology, particularly in flood management, which can fulfil all the requirements for flood prediction, preparation, prevention, and damage assessment (Tehrany, Pradhan, & Jebur 2013). Among different GIS-based flood models presented in the literature, artificial neural networks (Kia et al., 2011), frequency ratio (FR) (Lee et al. 2012), logistic regression (Pradhan 2010), adaptive network-based fuzzy inference system (Chau et al. 2005), multi-layered feed forward network (Kar et al., 2015), decision trees (Tingsanchali & Karim 2010; Merz et al. 2013; Tehrany et al. 2013), and support vector machines (SVMs) (Zhou et al. 2013; Tehrany et al. 2014) are the most widespread techniques that utilize RS and GIS tools. Although flood forecasting and prediction models are available, the accuracy of flood prediction maps remains a critical issue. In flood modelling, a high accuracy for flood prediction mapping should be achieved, and thus, new and efficient models should be explored to increase the accuracy.

Flood risk can be expressed as a combination of hazard and vulnerability (Apel et al., 2008; Intergovernmental Panel on Climate Change 2014; Vojinović & Abbott 2012). In particular, risk is a mathematical expectation of the vulnerability (consequence) function. Flood probabilities are determined to produce flood hazard maps. Hydraulic models may result in uncertainties because they require complete and sufficient hydrological data (Horritt 2006); therefore, using RS data and GIS-based models can be considered a complementary approach to flood modelling (Lecca et al. 2011).

The current research aims to determine the flood risk level in a study area that is affected by annual flooding. An appropriate flood prediction assessment is required in the study area to prevent incurring additional cost in the future because of the proximity of the area to an expressway, whose construction has already reached an exorbitant cost. As the main components of risk assessment, flood hazard and vulnerability indices were developed in this research through an accurate RS data and GIS-based ensemble model. The result of this research will provide an overall and accurate picture of flood vulnerable areas using detailed information to protect lives and properties along the New Klang Valley Expressway (NKVE) in case of flood events by implementing appropriate disaster management techniques.

2. Study area

Severe flood events that occurred during the past decades in Malaysia have seriously threatened its population, economy, and environment. This finding is evident from the increase in the amount of damages caused by a series of extreme floods in Malaysia over the last 50 years (Tehrany et al. 2013). One of the huge economic disadvantages of flooding is damages to highways. Thus, the Damansara River catchment along NKVE, which is seriously affected by flooding, was selected as the study area for a detailed flood hazard and risk analysis and modelling. The Damansara River catchment is located beside Kampung Baru Subang in Selangor, Malaysia. The study area is situated at 3°8'45.6" latitude and 101°32'27.24" longitude. Damansara River stretches from Sungai Buloh to Shah Alam. The location and boundary of the study area were accurately delineated using a digital elevation model (DEM), as shown in Figure 1.

As shown in Table 1, information regarding the hydro-geomorphological characteristics of the study area, such as basin slope, area, and length, was calculated using GIS Spatial Analyst tools. First, a flow direction map was extracted from the DEM from which the watershed boundary was delineated using hydrology basin tool. Next, main river length and upstream distance were calculated by using the flow length option in hydrology toolbox. Next, basin slope was calculated by using surface raster-based tool in percentage. Then using Zonal statistical analysis, average value was extracted. The total area of the Damansara River catchment was estimated at approximately 116.9 km², while the length of Damansara River was 22.22 km. Figure 2 shows the Damansara River catchment with its flow direction over the study area.

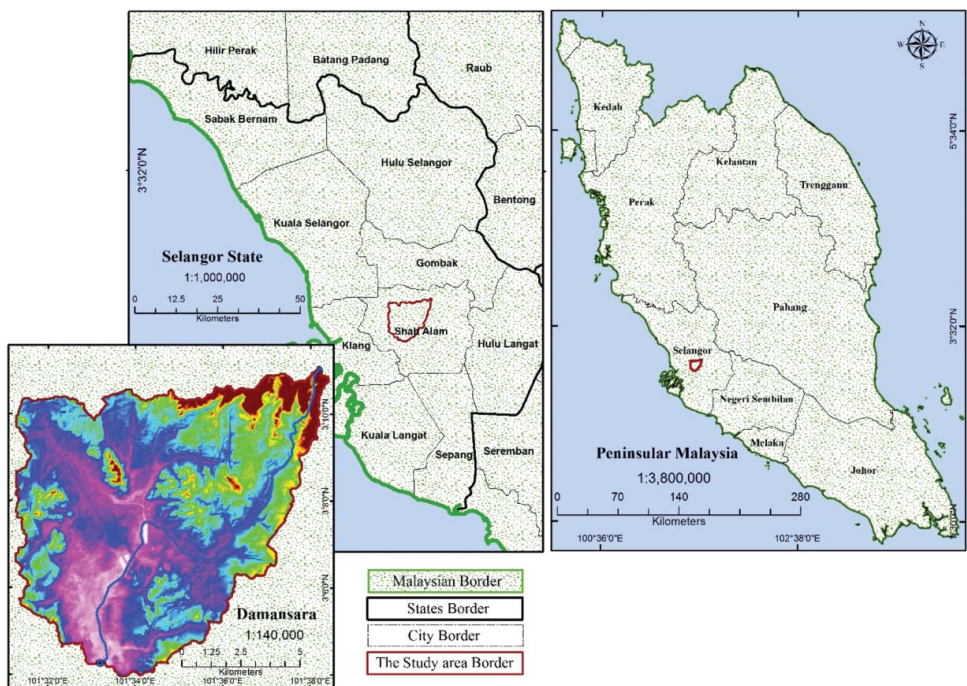


Figure 1. Location map of the study area.

3. Data used

3.1. Flood inventory

To evaluate flood risk in an area, analysing records of past flood events is essential (Manandhar 2010). Therefore, an inventory map is considered the most essential factor for predicting future disaster occurrence; such map can represent single or multiple events in a specific area (Tien Bui et al. 2012a). In the current research, a flood inventory map was primarily created by mapping the single flood locations where the exceed water has been running by using the field measurement and surveying. The prepared flood inventory map comprised 110 fluvial-flooded events, when collected from 2010 to 2015 over Damansara river catchment.

Table 1. Hydro-geomorphological details for Damansara river catchment.

Hydrological characteristics	Value
Basin slope (D)	0.07
Maximum flow distance (m)	22222.66
Max flow distance slope (D)	0.01
Stream centroid to outlet (m)	8426.41
Area facing south	53%
Area facing north	47%
Maximum stream length (m)	21473.58
Maximum stream slope (D)	0.00
Basin length (m)	15625.50
Basin shape factor	2.09
Basin perimeter (m)	71440.35
Basin average elevation (m)	37.86
CENTROID X	397595.17
CENTROID Y	346340.79
Basin LA county lateral	65.00
Basin LA county basin hydrograph	2.00
Basin area (km ²)	116.90

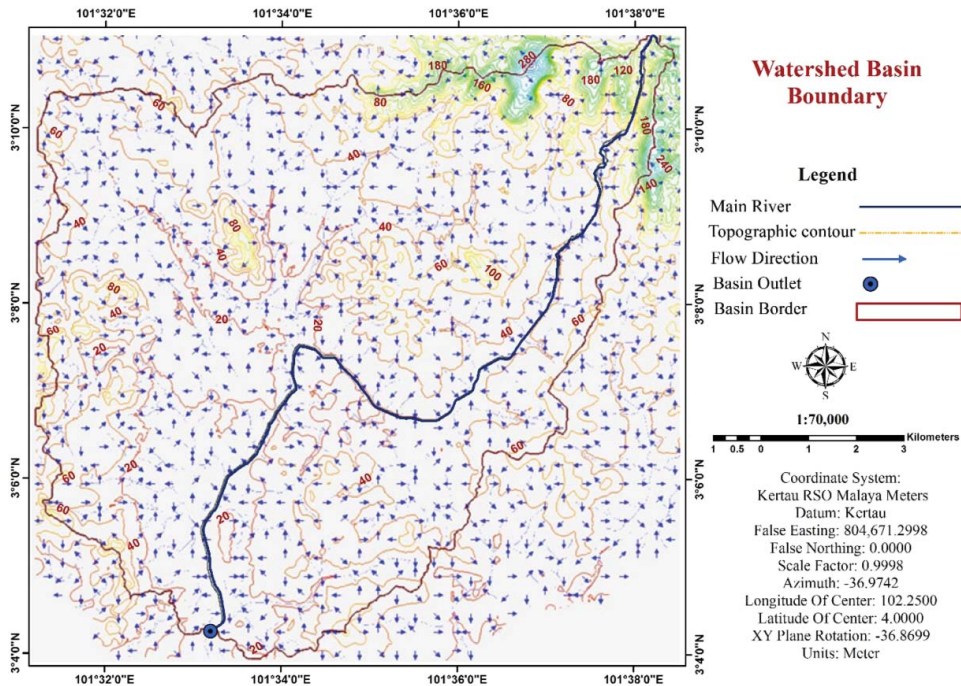


Figure 2. Watershed basin area with flow direction of Damansara river catchment.

The flood inventory map was divided into 70% training area and 30% validation area (Tunusluoglu et al., 2007), as shown in Figure 3. The training flood locations (77 out of 110 points) were randomly selected. Then, the flood layer, which was considered a dependent factor, was constructed. The flood layer was developed using two sets of value, namely, 0 and 1. 0 indicates the absence of flood events, whereas 1 indicates the presence of flood events across an area. Similarly, an equal number of points (77 out of 110) were selected as non-flooded areas to achieve a value of 0. Flooding could not occur in high-elevation regions such as hills; hence, non-flooded areas were randomly selected from these locations. The remaining flood locations (33 points) were utilized for model validation.

3.2. Flood-conditioning factors

In recent years, many spatial methods have been proposed by researchers for mapping flood hazard and risk zones to spatially delineate flood-prone areas. Building a flood hazard assessment model requires a set of flood-related parameters (Tehrany et al. 2014). The precision and quality of methods can be affected by the manner in which an accurate GIS database is used. Therefore, flood-conditioning factors should be optimized to enhance results.

The flood-conditioning factor data set used in this research consisted of 13 factors, namely, altitude, aspect, slope, curvature, stream power index (SPI), topographic wetness index (TWI), sediment transport index (STI), topographic roughness index (TRI), distance from river, soil, geology, surface runoff, and land use/cover (LULC). In this research, the cell of each conditioning factor was resized to $5\text{ m} \times 5\text{ m}$, and the grid of the Damansara River catchment was constructed with 2650 columns and 2623 rows.

3.2.1. DEM-derived factors

The DEM was built using Interferometric Synthetic Aperture Radar (IFSAR) images with a pixel size of $5\text{ m} \times 5\text{ m}$ which was captured in 2014. Consequently, all topographical map factors, such as altitude, slope, aspect, curvature, STI, SPI, TWI, and TRI, were derived from the DEM.

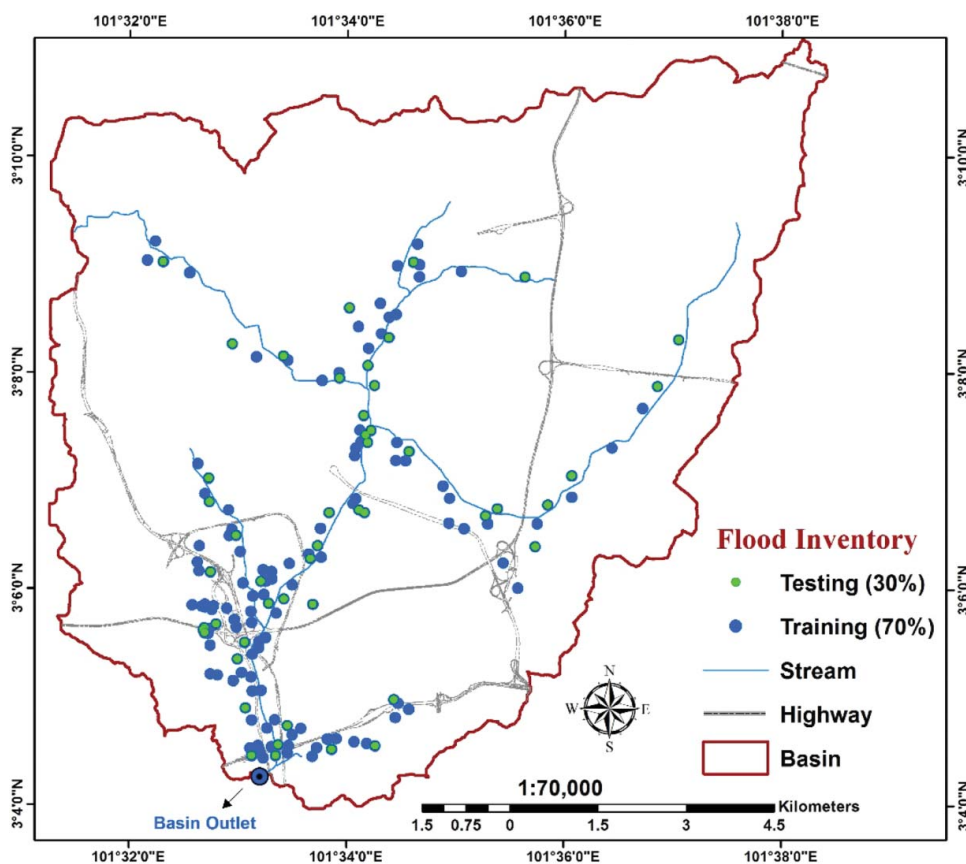


Figure 3. Flood inventory map.

One of the most influential parameters in flood studies is elevation (Figure 4(a)), and a flood event occurring in highly elevated areas is nearly impossible (Botzen et al. 2012). Water flows from highly elevated terrains toward lower regions, and thus, the probability of flood occurrence is naturally higher in flat regions. Moreover, topographical parameters that are directly affected by flow extent and runoff speed have important roles in flood occurrence (Kia et al. 2011). Each topographical parameter related to flood occurrence in any area is extracted directly from the DEM. Thus, a highly precise DEM is essential (Pradhan 2009).

Slope is another topographical factor that is regarded as an important parameter in hydrology (Tehrany et al. 2013) because of its effect in producing runoff in an area and its influence on runoff speed. An increase in slope degree decreases time for surface infiltration; subsequently, a huge amount of water enters the drainage network and causes a flood event (Figure 4(b)). Slope aspect commonly refers to the horizontal direction toward which the slope of a mountain is facing. An aspect map also plays a significant role in assessing the slope stability of a local terrain, depending on the type of slope face (Figure 4(c)). Curvature, which is split into three classes (convex, concave, and flat regions), is another influential parameter in flood occurrence (Figure 4(d)).

SPI and TWI are water-related parameters that are calculated using the following formulas (Gokceoglu et al. 2005):

$$\text{SPI} = \text{As} \tan \beta, \quad (1)$$

$$\text{TWI} = \ln(\text{As}/\tan \beta), \quad (2)$$

where A_s is the specific catchment area (m^2m^{-1}), and β is the local slope gradient measured in degree.

The SPI factor indicates the erosive power of water flow (Figure 4(e)). TWI represents the effect of topography on runoff generation and the amount of flow accumulation at any location in the river catchment (Gokceoglu et al. 2005), as shown in Figure 4(f).

The accuracy of a topographic index can be estimated with regard to grid spacing and terrain roughness by comparing the topographic index surface with respect to reference data. TRI, as one of the morphological parameters widely used in flood analysis, is calculated using the following equation:

$$\text{TRI} = \sqrt{\text{Abs}(\max^2 - \min^2)}, \quad (3)$$

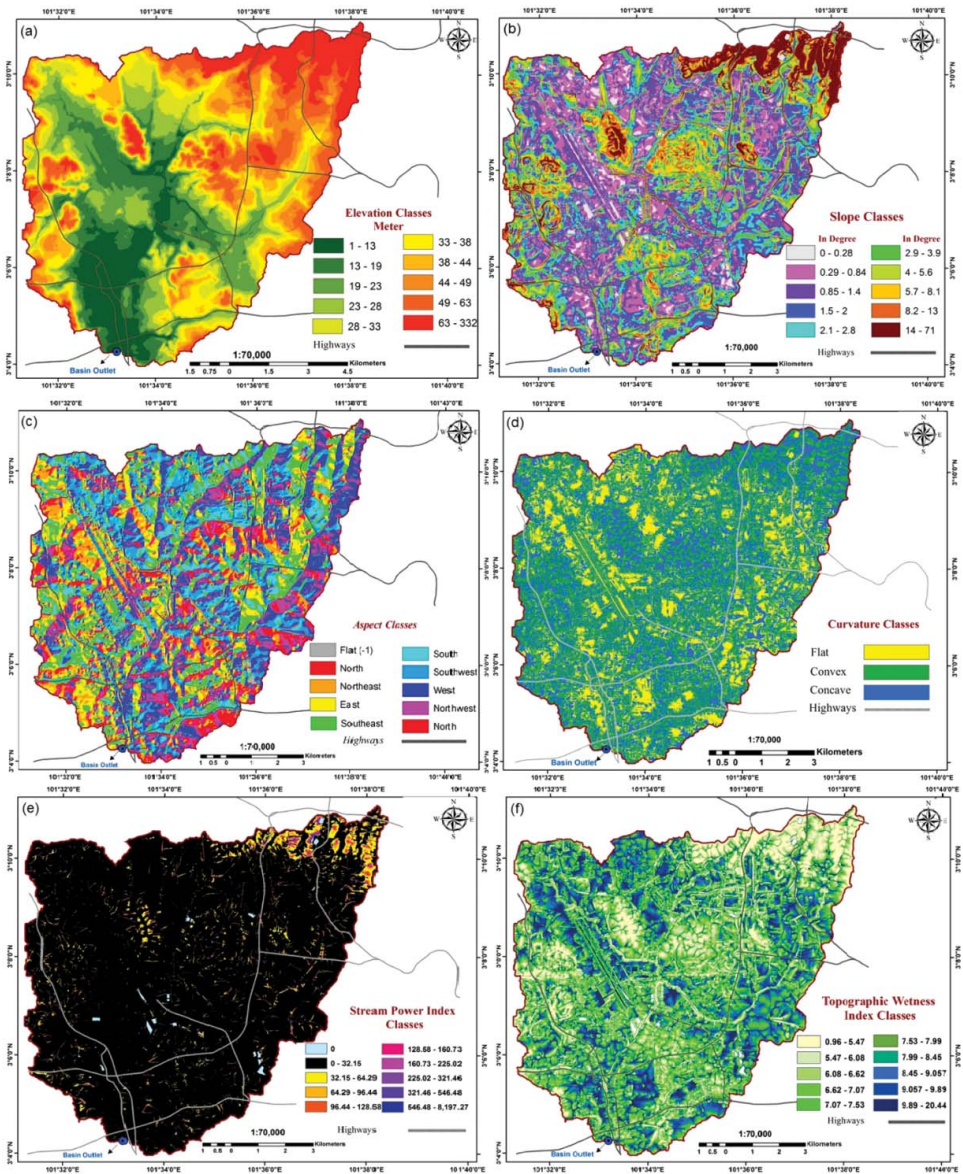


Figure 4. Flood-conditioning factors: (a) altitude, (b) slope, (c) aspect, (d) curvature, (e) SPI, (f) TWI, (g) TRI, (h) STI, (i) distance from river, (j) soil, (k) geology, and (l) surface runoff.

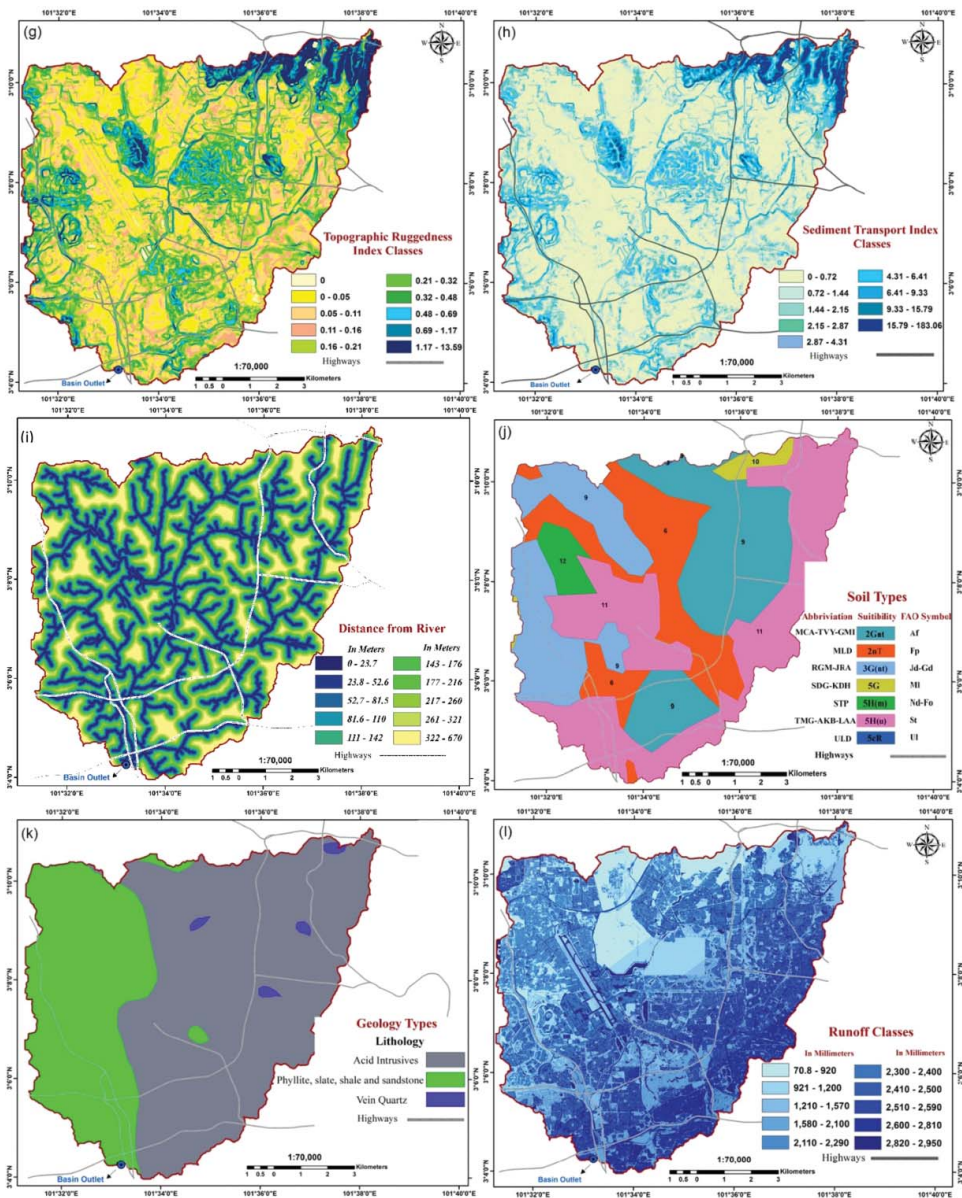


Figure 4. (Continued)

where max and min are largest and smallest values of the cells in the nine rectangular neighbourhoods of altitude (Figure 4(g)).

The erosion and deposition processes are characterized using STI (Figure 4(h)), as presented in the following equation (Moore & Wilson 1992):

$$STI = (m+1) \times (As/22.13)^m \times \sin(B/0.0896)^n, \quad (4)$$

where As is the specific catchment area (i.e., the upslope contributing area per unit contour length) estimated using one of the available flow accumulation algorithms in the Hydrology toolbox of

ArcGIS, and B is the local slope gradient in degrees. The contributing area exponent, m , is generally set to 0.4, whereas the slope exponent, n , is generally set to 1.4.

3.2.2. Distance from river

Flood occurrences in the study area are frequent along the stream. Thus, distance from the river was considered another geomorphology-related conditioning factor. Subsequently, a distance from the river map was generated because the streams would disrupt the stability of the slopes either by toe undercutting or by saturating parts of the materials lying within the water level of stream ways. Distance from a river is represented by the proximity of rivers and drainages in an area. In the current research, a distance from the river map was developed from the vector map of rivers using Euclidean distance in ArcGIS 10.3 software. Then, the resulting shapefile was converted into a 5 m raster and divided into 10 classes using the quantile method (Figure 4(i)).

3.2.3. Lithological and soil type

Lithological and soil maps are highly important parameters in finding sensitive areas prone to flooding. Soil type directly affects the drainage process because of soil characteristics, such as texture, permeability degree, and structure. The study area is characterized by seven soil types. The spatial distribution of each soil type is shown in Figure 4(j). Lithological information regarding the permeability of rocks is also required in flood hazard assessment. Therefore, soil types and lithology are vital for conducting analysis in this research. Figure 4(k) presents the geology map, which shows that three types of lithology cover the study area. The majority of the eastern section of the study area is covered by acid intrusives, whereas the western section is covered by phyllite, slate, shale, and sandstone.

3.2.4. Surface runoff

Soil capacity is fully saturated by water throughout the land, and water flow exceeds the limits required for surface runoff (Figure 4(l)). This parameter was estimated using an empirical equation called the Soil Conservation Service curve number (SCS-CN) method,

$$Q = \frac{(P - 0.2s)}{P + 0.8s} \quad (5)$$

$$S = \frac{1000}{CN} - 10, \quad (6)$$

where Q is the direct runoff (mm), P is the accumulated rainfall (mm), S is the potential maximum soil retention (mm), and CN is the curve number.

3.2.5. LULC

Another primary related factor that strongly contributes to flooding is LULC. A detailed understanding of LULC is extremely essential for environmental and natural hazards (Rizeei et al. 2016). Vegetated areas are less prone to flooding because of the negative correlation between a flood event and vegetation density. However, urban areas are typically composed of impermeable surfaces and bare lands, which increase storm water runoff. In this research, a land-use map played a crucial role in flood hazard modelling as one of the conditioning factors and criteria for vulnerability assessment. Therefore, considering the importance of this factor, a very high resolution image obtained from the WorldView-3 satellite was used to extract an LULC map. The WorldView-3 satellite is the first multi-payload, super-spectral, high-resolution commercial satellite that operates at an altitude of 617 km. It provides 31 cm panchromatic resolution, 1.24 m multi-spectral resolution, 3.7 m short wave infrared resolution, and 30 m CAVIS resolution, that was captured on 9th of December 2014.

After performing preprocessing approaches, such as geometric, radiometric, and atmospheric corrections, the object-based SVM algorithm was implemented using ENVI 5.3 to classify the

Table 2. Details of SVM segmentation and classification method.

Segment setting		Merge setting		Textural Kernel size	Classification algorithm				
Scale level	Algorithm	Merge level	Algorithm		Method	Threshold	Type of kernel	Gamma in kernel	Penalty parameter
50	Edge	90	Full lambda schedule	3	SVM	5	Radial basis	0.03	100

Worldview-3 image. The details of the SVM segmentation and classification method are provided in Table 2.

The final land-use map presented in Figure 5 classifies land use into seven classes, namely, high-way, bare land, forest, built-up land, green and recreation areas, road, and water body. The study area is mostly covered by built-up land and forest.

4. Methodology

The methodology applied in the present research comprises different phases that are illustrated in the overall flowchart shown in Figure 6. As defined, flood risk is generally represented as the product of a hazard and the vulnerability of an exposed environment (Müller et al. 2011). The spatial hazard model was generated in GIS environment using a combination of a probability map obtained from an ensemble model and rainfall as a triggering factor. The risk map was also produced by integrating hazard and vulnerability maps to show the flood risk levels across the study area. To apply the GIS-based ensemble model for flood hazard modelling, each conditioning factor was classified using the quantile method as a requirement of FR modelling (Ayalew & Yamagishi 2005). Then, the FR model was applied, and each FR value was assigned to each class of conditioning factor. Each flood-related layer was built in ArcGIS and then transformed into ASCII format and entered as input in the SPSS modeller for SVM analysis.

4.1. Optimizing flood-conditioning factors using FR

Flood-conditioning factors should be identified to evaluate flood probability throughout a specific time and in a particular environment (Yalcin et al. 2011). Applying the FR method as one of the GIS-based approaches can considerably contribute in identifying the effect of each flood-related parameter on flood events in a study area (Tehrany et al. 2013). The FR value illustrates the relationship between each class of conditioning factor and flood location, so weights will be precisely assigned to each class under each factor (Neshat & Pradhan 2015). An FR value > 1 indicates a strong relationship, whereas an FR value < 1 denotes a weak correlation (Akgun et al. 2007), as shown in Table 3.

The estimated FR ratio for weighting each conditioning factor was normalized within the range of 0–1. Each conditioning factor varies in dimension, and thus, normalization should be performed to make a factor appropriate for use as a direct input for SVM modelling. A popular technique used for the normalization process is as follows (Choi et al., 2009):

$$Y_i = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \quad (7)$$

If y_i ($i = 1, 2, \dots, n$), then Y_i indicates the normalized values of y_i . y_{\min} and y_{\max} represent the minimum and maximum value of y_i , respectively. Consequently, each data category, including the nominal and interval classes, was transformed into a single scale ranging from 0 to 1 to enter as input into the SVM model.

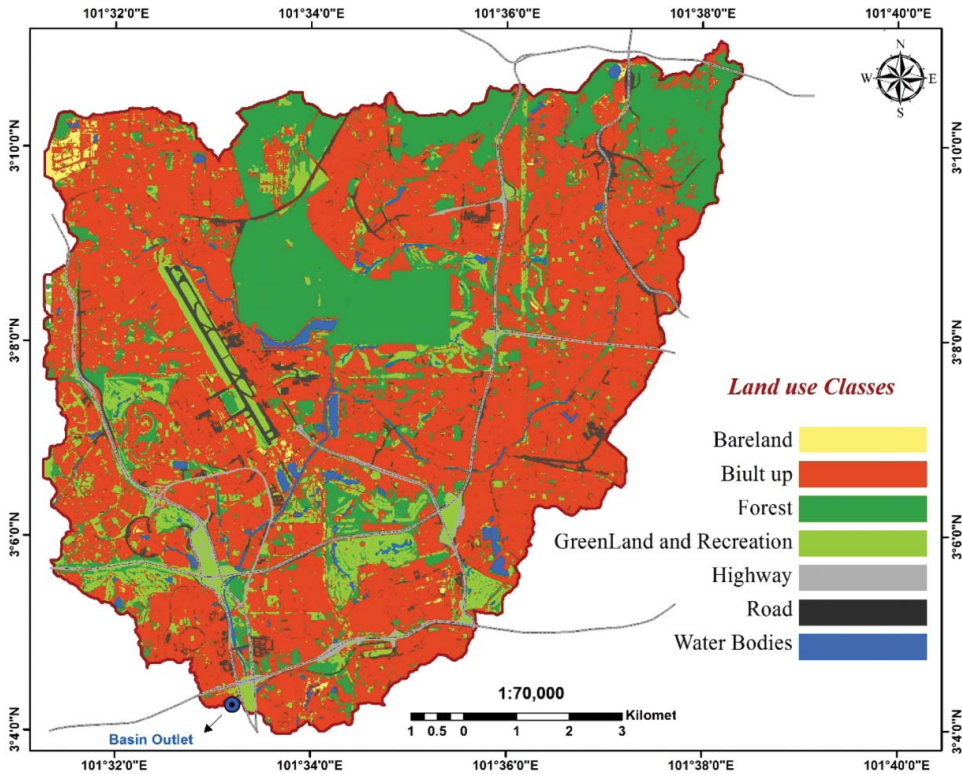


Figure 5. Classified land-use map used as a flood-conditioning factor and vulnerability criteria assessment.

4.2. Flood probability evaluation using SVM

SVM is based on statistical learning theory to minimize operational risk standard (Yao et al. 2008). Nonlinear structures can be converted into linear structures because of the creation of a hyperplane, which can generate the process (Jebur et al. 2014a). The transformation of data via a mathematical function is identified as the kernel function. The basis of this method is the hyperplane formation separation of the training data set. A separating hyperplane is created in the original space of n coordinates (x_i parameters in vector x) between the points of two distinct classes (Marjanović et al. 2011).

The peak edge of separation is found among the classes via SVM, and the hyperplane is classified in the central part of the peak edge (Marjanović et al. 2011). The point classification will be based on hyperplane changes, and is classified as 1 if it will be overhead the hyper-plane, and as -1 if it will not be overhead. Through this rule, the new data feature can be used to anticipate the set to which a different record should fit. Support vectors are known as the neighbouring training points of the optimal hyperplane.

For example, assume a training data set of instance-label pairs (x_i, y_i) with $x_i \in R^n$, $y_i \in \{1, -1\}$, and $i = 1, \dots, m$. In the current flood probability estimation case, x is a vector of each input space, including altitude, aspect, slope, curvature, TWI, SPI, TRI, STI, distance from river, lithology, soil, surface runoff, and land use. Both flooded and non-flooded pixels are illustrated using two classes $\{1, -1\}$. Thus, recognizing the optimal separating hyperplane is the objective of the SVM model. For the case of linear separable data, a separating hyperplane can be defined as follows:

$$y_i(w \cdot x_i + b) \geq 1 - \zeta_i \quad (8)$$

where w is the norm of the normal of the hyperplane, b is the offset of the hyperplane from the origin, and ζ_i denotes positive slack variables. The optimization problem can be solved by designating

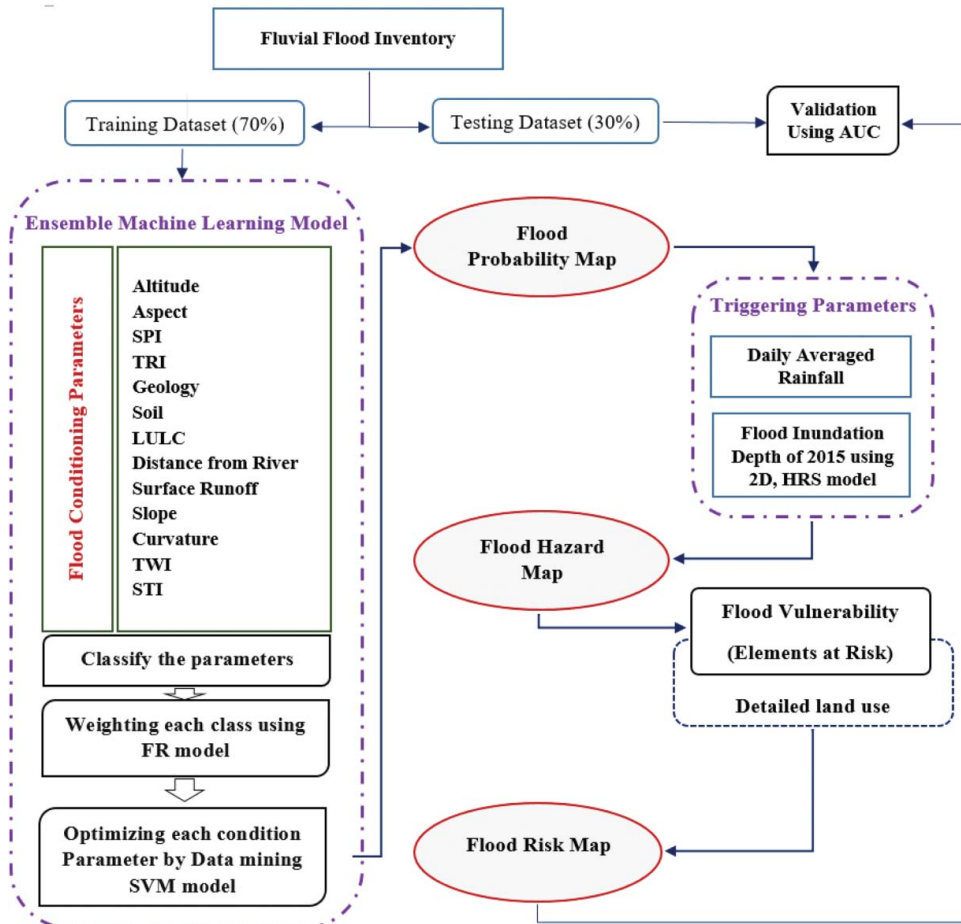


Figure 6. Overall methodology flowchart.

an optimal hyperplane using Lagrangian multipliers (Samui 2008).

$$\text{Minimize } \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (x_i x_j) \quad (9)$$

$$\text{Subject to } \sum_{i=1}^n \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq c \quad (10)$$

where α_i denotes the Lagrange multipliers, C is the penalty, and the slack variables ζ_i allow penalized constraint violation.

When the hyperplane is not separated using the linear kernel function, the initial input data may be transformed into a high-dimensional feature space using several nonlinear kernel functions. New data classification will be performed through the decision function, which is described as follows:

$$g(x) = \text{sign} \left(\sum_{i=1}^n y_i \alpha_j K(x_i, x_j) + b \right) \quad (11)$$

where $K(x_i, x_j)$ is the kernel function.

Table 3. FR spatial correlation between flooded area and each conditioning factor.

Factor	Class	FR	Factor	Class
Elevation (m)	1–13	2.98	TWI	0.96–5.47
	13–19	2.34		5.47–6.08
	19–23	2.08		6.08–6.62
	23–28	1.49		6.62–7.07
	28–33	1.08		7.07–7.53
	33–38	0.89		7.53–7.99
	38–44	0.68		7.99–8.45
	44–49	0.64		8.45–9.057
	49–63	0		9.057–9.89
Slope (degree)	63–332	0	STI	9.89–20.44
	0–1.68	1.7		0
	1.69–4.48	1.35		0–0.72
	4.49–7.84	1.03		0.72–1.44
	7.85–11.8	1.28		1.44–2.15
	11.9–16.2	1.17		2.15–2.87
	16.3–21	1.09		2.87–4.31
	21.1–26.3	1.01		4.31–6.41
	26.4–33	0.72		6.41–9.33
	33.1–45.6	0.6		9.33–15.79
	45.7–71.4	0.06		15.79–183.01
Aspect	Flat	1.5	TRI	0
	North	1.04		0–0.05
	Northeast	1.05		0.05–0.11
	East	0.99		0.11–0.16
	Southeast	1.59		0.16–0.21
	South	1.56		0.21–0.32
	Southwest	0.57		0.32–0.48
	West	0.35		0.48–0.69
	Northwest	0.59		0.69–1.17
				1.17–13.59
Land use	Bare land	0.69	Soil type	1 (TMG-AKB-LAA)
	Forest	0.24		2 (MCA-TVY-GMI)
	Highway	0.16		3 (RGM-JRA)
	Recreation	2.7		4 (SDG-KDH)
	Road	0.66		5 (STP)
	Urban	0.77		6 (ULD)
	Water	0.42		7 (MLD)
Geology	Acid intrusive	0.8	Curvature	Concave
	Phyllite, slate, shale and sandstone	0.93		Flat
	Vein quartz			Convex
SPI	0	0	Distance from river (m)	0–62
	0–32.15	1.05		62–65
	32.15–64.29	0.37		65–67
	64.29–96.44	0.31		67–90
	96.44–128.58	0.26		90–96
	128.58–160.73	0.21		96–110
	160.73–225.02	0.2		110–115
	225.02–321.46	0.15		115–122
	321.46–546.48	0.12		122–124
	546.48–8,197.27	0		124–143
Surface runoff	81.2–177	0.19		
	178–518	0.51		
	519–582	1.1		
	583–603	1.9		
	604–624	1		
	625–1020	1.17		
	1030–1490	1.35		
	1500–1750	1.74		
	1760–1820	1.84		
	1830–2800	1.68		

Kernel type selection in an SVM model can be considered a vital step because it directly controls effective training and classification accuracy (Yao et al. 2008). Linear (LN), polynomial (PL), radial basis function (RBF), and sigmoid (SIG) are the four kernel types used in SVM. LN is regarded as a distinctive case of RBF although SIG execution is equivalent to that of RBF for the given factors (Song et al. 2011). When RBF is used in the processing, LN is no longer required. In terms of accuracy, RBF generates more reliable and solid outcomes compared with SIG because of its higher fitness in interpolation. A potential shortcoming of RBF is its failure to create long-range extrapolation. By contrast, PL exhibits better extrapolation fitness (Tehrany et al. 2014). RBF was used in the present study to estimate flood occurrence probability.

4.3. Flood risk evaluation

Risk analysis mainly aims to determine the probability of a specific hazard that will result in damage. The correlation between the frequency of a disastrous event and the intensity of its consequences is determined via risk evaluation. In the present study, flood risk evaluation aims to ascertain the expected degree of loss because of a flood event. The 'risk' (R) is commonly expressed as follows:

$$R = f(H_L, V_L) \quad (12)$$

where H_L and V_L represents flood hazard and flood vulnerability, respectively.

As discussed earlier, a hazard map is one of the essential components of flood risk analysis. The assumption that rainfall is one of the primary triggering factors of flood occurrence over a study area, which results in extreme events such as flooding and overflowing, contributes to the preparation of a flood hazard map. The average daily Rainfall data were obtained from 15 rainfall gauge stations in and around the study area. Then, the daily average precipitation for 6 years (2010 to 2015) was used to create the rainfall density map with the inverse distance weighting (IDW) interpolation model. Another hazardous triggering factor is 'Flood inundation depth' which can be estimated by using 2D HRS model. 2D high-resolution sub-grid models take an advantage of wetting and drying algorithm. Mostly in numerical flood models for surface flows, drying and wetting algorithm is basically reined by artificially placing which called screens in grid's velocity points once the water depth drops fall below a defined drying threshold, while eliminating the Screens once the water depth rises beyond a flooding threshold (Casulli 2009).

The flow hydrograph is calculated to bring a stream flow into the 2D flow area. There are some requirement for this analysis: (1) flow hydrograph (Q/t), and (2) energy slope of stream (degree).

In this research hourly stream, flow has been used since February 2010 to February 2015 in order to model the maximum flood inundation depth. For each river reach a related station was used to generate stream flow and then input to the model for unsteady analysis. Stream flow data recorded at four gauging stations have been used as the upstream boundary condition. Figure 7 presents the location of water level, rainfall and rain gauges station in and around the Damansara river catchment.

Maximum flood inundation depth then integrates with daily average of precipitation in order to create the hazardous triggering layer. The flood probability map is transformed into a hazard map by multiplying it with hazardous triggering layer. To calculate the hazard map, the following equation is used:

$$H = f(P_S, T_{pi}) \quad (13)$$

where H indicates the hazard probability, P_S indicates the probability obtained from the ensemble of FR and SVM analysis, and T_{pi} is the hazardous triggering layer. Moreover, these two maps should be the standardized into a common dimensionless scale before they are combined given that the scales of

their data are different from each other. The following equation is used for the standardization:

$$X_{ij} = \frac{X_j - X_{ij}}{X_{\max-j} - X_{\min-j}} \quad (14)$$

where X_{ij} represents the standardized score for the i th alternative and the j th attribute; X_{ij} is the raw score; and $X_{\max-j}$ and $X_{\min-j}$ are the maximum and minimum scores for the j th attribute, respectively.

Vulnerability is another indispensable factor in flood risk evaluation; it is commonly regarded as a factor that leads to circumstances that make a system or an individual prone to damage caused by a hazard (Muller, Reiter, & Weiland 2011). A set of site-specific parameters is frequently used to evaluate vulnerability.

Flood vulnerability (V_L) can be defined mathematically as follows:

$$V_L = P[D_L \geq 0|L], \quad (0 \leq D_L \leq 1) \quad (15)$$

where D_L is the assessed (definite) or the expected (forecasted) damage to an element given the occurrence of a hazardous flood event (L). Vulnerability is the probability of total loss to a specific element or the proportion of damage to an element given the occurrence of a flood event. In both cases, vulnerability is expressed on a scale from 0 to 5, with 0 indicating the lack of damage and 5 indicating complete loss or destruction. Vulnerability to floods is expressed in economic (monetary, quantitative) and heuristic (qualitative) scales. When using economic measures, vulnerability is most commonly expressed in terms of element value, such as monetary, intrinsic, and utilitarian values. When expressed heuristically, flood vulnerability is described in a qualitative (descriptive) term that indicates the expected or definite damage to an element at risk.

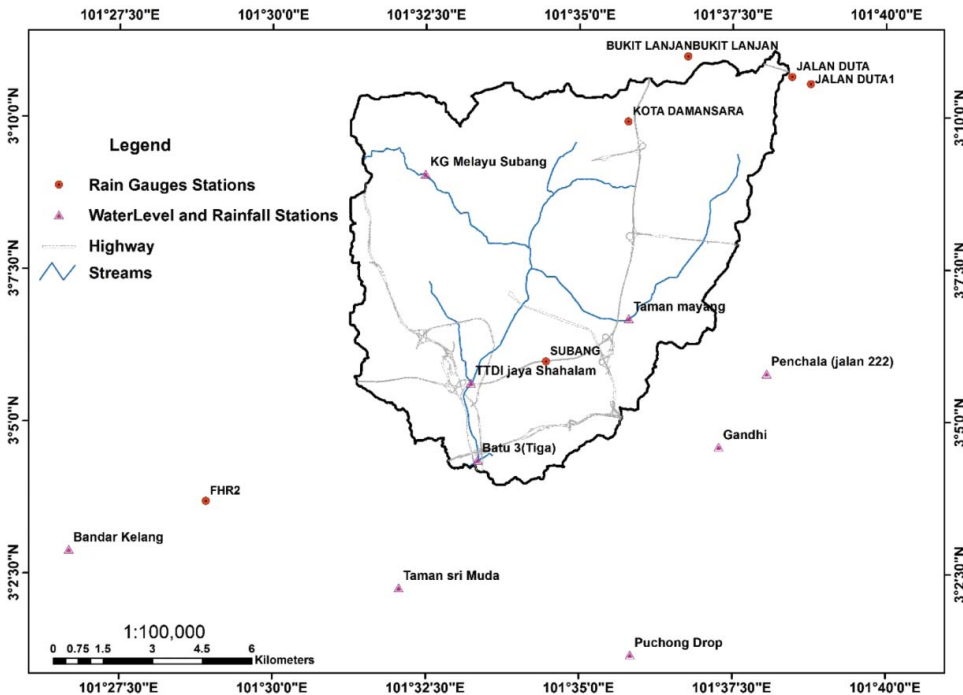


Figure 7. Location of water-level, rainfall and rain gauges station in and around the Damansara river catchment.

4.4. Validation

Validation is another important process that should be performed to evaluate the efficiency and precision of the derived result. The area under the curve (AUC) method has been widely used in numerous studies to estimate the performance of probability modelling (Althuwaynee et al. 2012). On the basis of this approach, the probability map is split into equal area classes and grades (from the least to the most value). The flood probability level can be specified through success and prediction curves. A curve is created by plotting the accumulative percentage of flood-prone areas (from the maximum to the minimum probability) on the x -axis and the accumulative percentage of flood locations on the y -axis. A steeper curve indicates a higher number of flood locations falling into the most-prone category. The AUC range varies from 0.5 to 1.0, and the highest accuracy has a value of 1.0, thereby suggesting that model performance is completely satisfactory in predicting disaster occurrence without any bias. Therefore, an AUC value that is closer to 1.0 indicates the precision and trustworthiness of the model.

5. Results and discussion

5.1. Flood probability map using the GIS-based ensemble method

As shown in Table 3, the correlation of each class of each conditioning factor was estimated by applying the FR model. The lowest class of altitude (ranging from 1 m to 13 m) obtains the highest FR value at 2.98, which is representative of the highest correlation of this class with flood occurrence. The natural behaviour of flooding, which occurs mostly in flat regions instead of in highly elevated areas, can provide an appropriate proof for these results. In the case of aspect, the class of southeast direction obtains the highest FR value at 1.59. The highest FR value is also assigned to slope range of 0–1.68 degree, which is at 1.70. In the case of the curvature factor, the flat areas obtain the highest correlation. The highest (1.10) followed by concave (0.93) expectedly. FR values are also assigned to the lowest (0–32.15) and highest (546.48–8197.27) values of the SPI factor, which means low power streams, are more prone to be flooded. For the TWI factor, the class range of 0.96–5.47 exhibits the lowest correlation (0.49) with flooding while class 9.89–20.44 has the highest correlation with flood which means more likely to be wet. In the case of the TRI and STI factors, the class ranges of 0–0.05 and 0–0.72 obtain the highest FR values at 1.46 and 1.44, respectively, which show in smooth area such as plain and alluvial type of river, the occurrence of flood is high and support the accuracy of this model.

Another significant parameter in flood occurrence is distance from river. From the results derived using the FR approach, the nearest class to river (0–62 m) obtains the highest FR ratio of 1.83. The second class of distance from river (62–65 m) also achieves a high value, which can be attributed to the long coverage of alluvial flood over the catchment.

In the case of geology, obviously, there was no meaningful correlation between its classes and flood events because all the classes assigned below one, although class of phyllite, slate, shale, and sandstone exerts more influence on flood occurrence rather than class of vein quartz. A strong relationship exists between each type of LULC and their effects on the speed of water flow and flood occurrence. In the study area, the regions covered by recreation areas achieve the highest FR value of 2.70. In the case of soil type, type 1 soil (TMG-AKB-LAA) has the highest FR value (2.34). However, soil types 2 (MCA-TVY-GMI) obtains the least value due to their soil components, which indicates that no significant correlation exists between these soil types and flooding. Lastly, the highest FR value (1.84) in the case of surface runoff is achieved at the range of 1760–1820 mm of runoff, it can be noticed that where we have high surface runoff is almost prone for flood as well.

When the FR value of each conditioning factor was evaluated, the weights of each range were normalized as shown in Equation (7). Then, each conditioning factor was reclassified based on the normalized FR weights. The reclassified conditioning factors were applied as input for SVM

modelling. As mentioned earlier, RBF kernel was used to generate a flood probability map in GIS environment. The probability index ranged from 0 to 1. [Figure 8](#) shows ensemble FR-SVM optimization result and [Figure 9](#) demonstrates the flood probability map generated from the ensemble FR-SVM model. The highest probability of flood occurrence is located in regions with the lowest elevation and direction of the basin outlet.

Based on the result of ensemble SVM and FR modelling which as shown in [Figure 8](#), geology, aspect, TRI and soil factors are achieved the lowest rank, respectively, and consider insignificant with flood probability in our study area. However, Altitude, Distance from River, STI and TWI are the most significant parameters which are related to flood occurrence, respectively. The correlation coefficient of this model was estimated as 0.7413 and mean absolute error was 0.2453.

Therefore, each class of conditioning factors along with each conditioning factor has been optimized by ensemble FR and SVM model. Then all weighted layers were overlaid to generate the flood probability map.

5.2. Flood risk mapping

In this study, a flood risk map was mainly derived from the combination of the hazard and vulnerability maps as discussed earlier. During the first phase of this research, the ensemble of the FR and SVM model was used to investigate flood probability across the catchment. During the second phase, after the probability map was obtained, rainfall and flood inundation depth were selected as the main triggering factor to combine with the probability of acquiring the hazard map as illustrated in [Equation \(13\)](#). A final hazard map ranging from 0 to 1 was obtained as a result of this combination. The highest hazard level occurs in the southern part of the study area, where the elevation toward the basin outlet is decreased. [Figure 10](#) shows the classified hazard map of the study area.

In a next phase, after the final hazard map was prepared, the vulnerability map was evaluated as a second initial component of the risk map. The vulnerability map was derived by relying on the data collected from GIS maps with information on detailed land-use types. As mentioned earlier, the land-use map was prepared using very high resolution satellite images from WorldView-3. The overall accuracy of classification was at 84.07%, which could be considered accurate as previously illustrated in [Table 3](#). The value of a criterion was assigned to each element at risk based on their

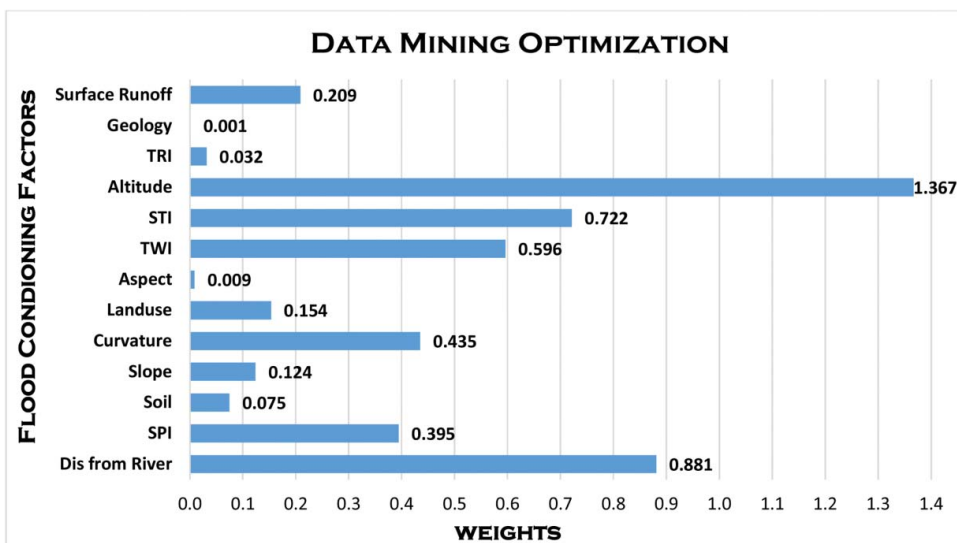


Figure 8. Ensemble machine-learning optimization result.

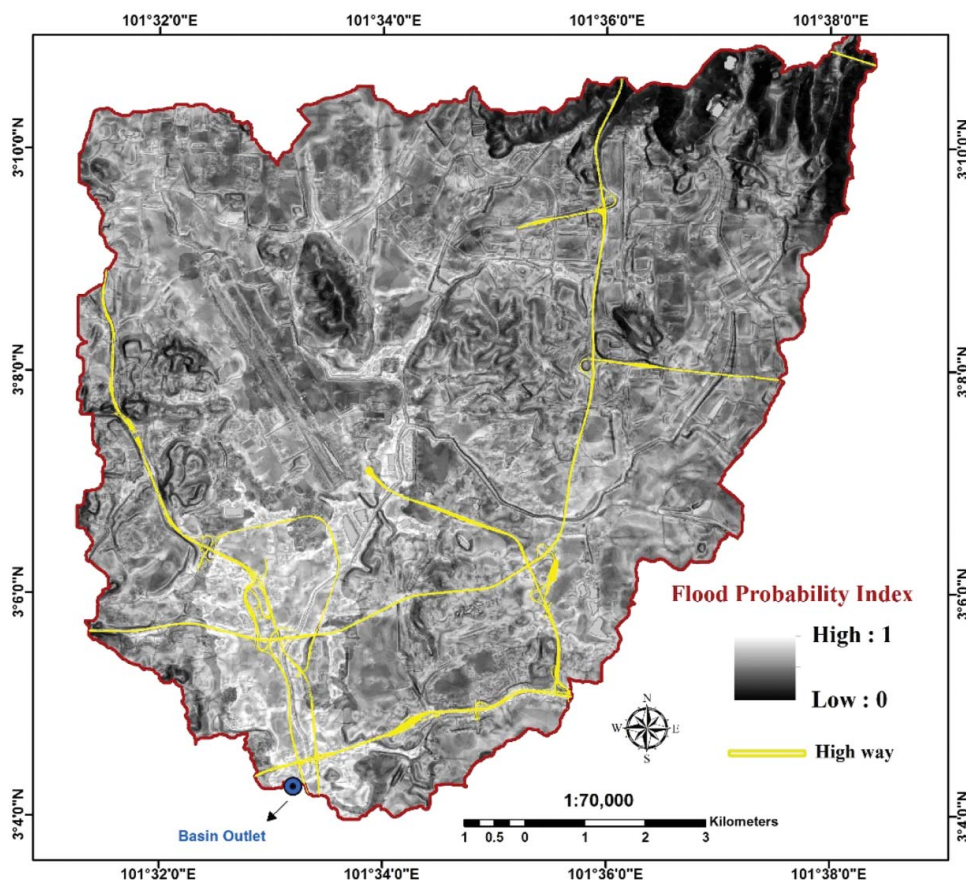


Figure 9. Flood probability map for Damansara river catchment.

importance in the study area. The majority of the catchment area is also covered by built-up land, where most of the residential areas and commercial centres are located. After the highway, vulnerability criteria value (5) was assigned to the built-up land. The second highest value of vulnerability (4) in this area was allocated to the highway because it is located at a lower elevation compared with the built-up land. In case of flood occurrence across the highway, loss of human lives and damage to properties will occur across the highway. Furthermore, the rescue operation of emergency workers, such as firefighters, police officers, and medical personnel, in the built-up land will be interrupted and delayed. Therefore, on the basis of expert opinion, the highway was determined as the second most vulnerable element at risk in the study area. The lower vulnerability value based on expert opinion was assigned to the road (3) and recreation area (2). However, the lowest vulnerability value (1) was allocated to areas covered by forest because they are densely vegetated, which reduces the speed of runoff.

The final stage of this research aims to determine flood risk in the study area based on the quantitative approach. Therefore, the flood hazard and vulnerability maps were combined to produce a direct specific risk map using Eq. 12, and a flood risk map ranging from 0 to 1 was achieved, then the final risk map was split into classified risk areas to improve visual interpretation. Hence, the flood risk levels were clearly recognized. When classifying the flood risk map into categorical classes, a quantile method was applied. This classifier is commonly recommended for this objective because it uses mean values to create class breaks. Five risk levels, namely, no risk, low, moderate, high, and

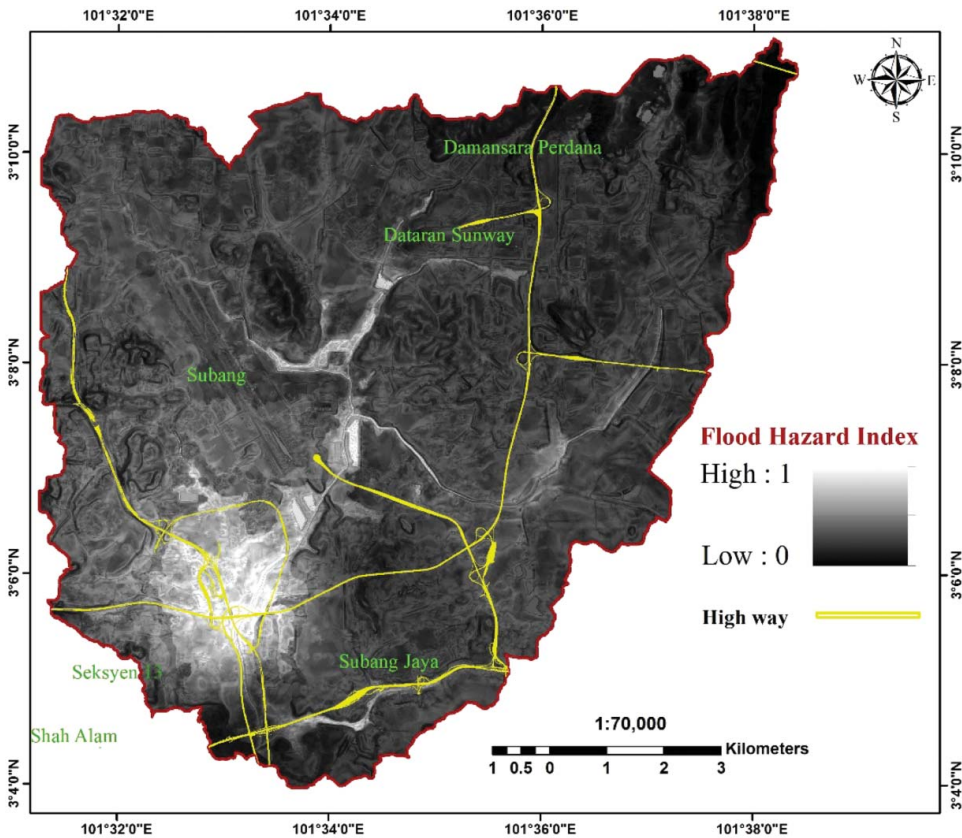


Figure 10. Flood hazard index map of Damansara river catchment.

very high, were distinguished using this approach (Figure 11). The percentages of each risk level in the study area are also presented in Table 4. No risk and low-risk areas cover 42.8% of the entire catchment. Moderate flood risk (30.2% of the catchment) is also depicted in regions close to the high- and very high risk zones in the study area. The high areas cover 22.6% and very high risk areas cover only 4.2% of the catchment.

5.3. Validation and field verification

The efficiency of the flood hazard map using the ensemble model was evaluated via receiver-operating characteristic (ROC) curve. The success rate and prediction rate curves were evaluated. The success rate and prediction rate curves were produced using the flood training data set (70%) and the flood validation data set (30%), respectively. The results are 89.7% and 78.9% for the success and prediction rates, respectively (Figure 12).

Furthermore, multiple field visits were conducted to verify the derived results (Figure 13). The flood location and condition was recorded during the fluvial flood events on 12 and 27 October 2014, 13 and 29 January 2015, and 18 February 2015. Expectedly, all the recorded inventory points fell within the high-risk and very high risk zones, where the probability of flood was very high. This process presents a good verification of the reliability of the applied model in the study area.

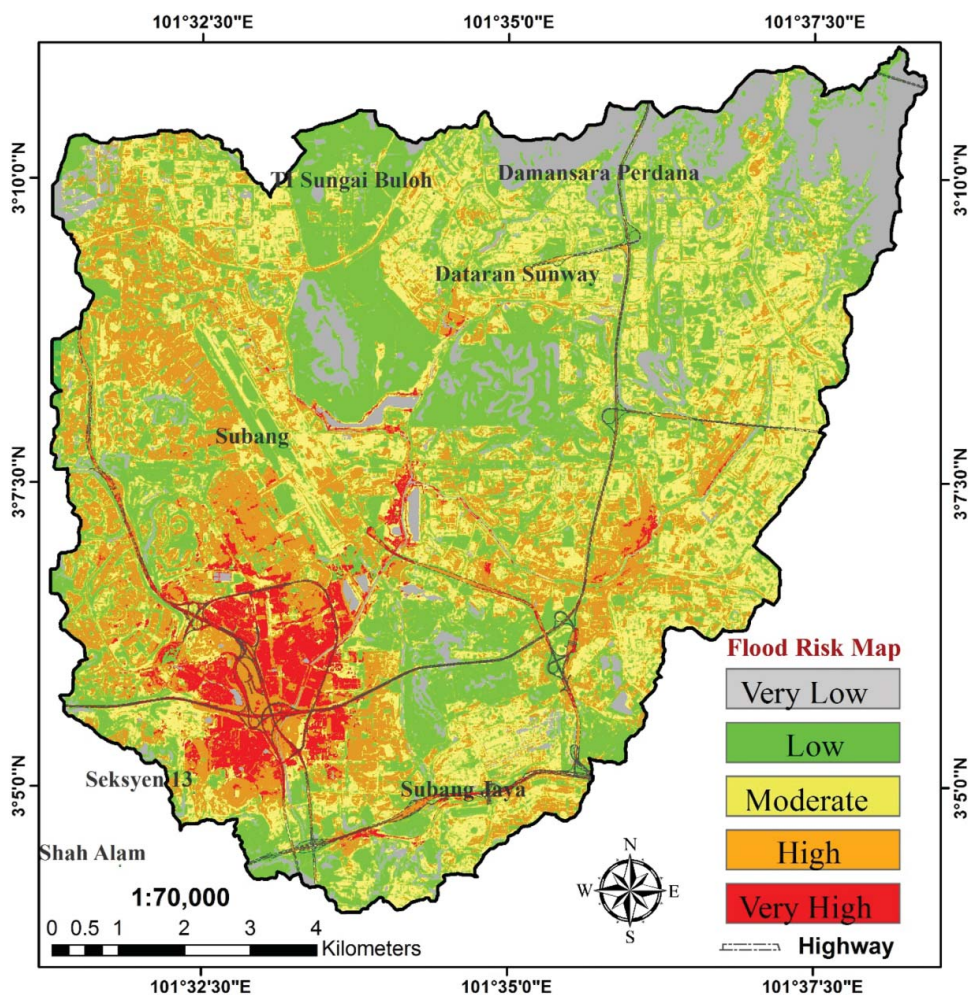


Figure 11. Flood risk levels of the study area.

6. Conclusion

Flood occurrence is a serious and disastrous event that can happen practically anywhere. Therefore, controlling the effects of flood is crucial, and can be performed via flood hazard and risk mapping. Flood-prone areas must be identified to anticipate and analyse the spatial distribution of appropriate flood management in the future. The research background indicates that various methods and techniques have been applied to identify flood-susceptible areas. In the current research, the GIS-based ensemble method of FR and SVM was used in the flood hazard mapping of the study area along

Table 4. The area percentage of risk map for Damansara river catchment.

Code	Class	Area (m ²)	Area (km ²)	Percentage
1	Very low	14900472.19	14.90	12.75
2	Low	35307217.61	35.31	30.21
3	Moderate	35329035.80	35.33	30.23
4	High	26440455.04	26.44	22.63
5	Very high	4880703.38	4.88	4.18
Total	116857884.03	116.86	100%	

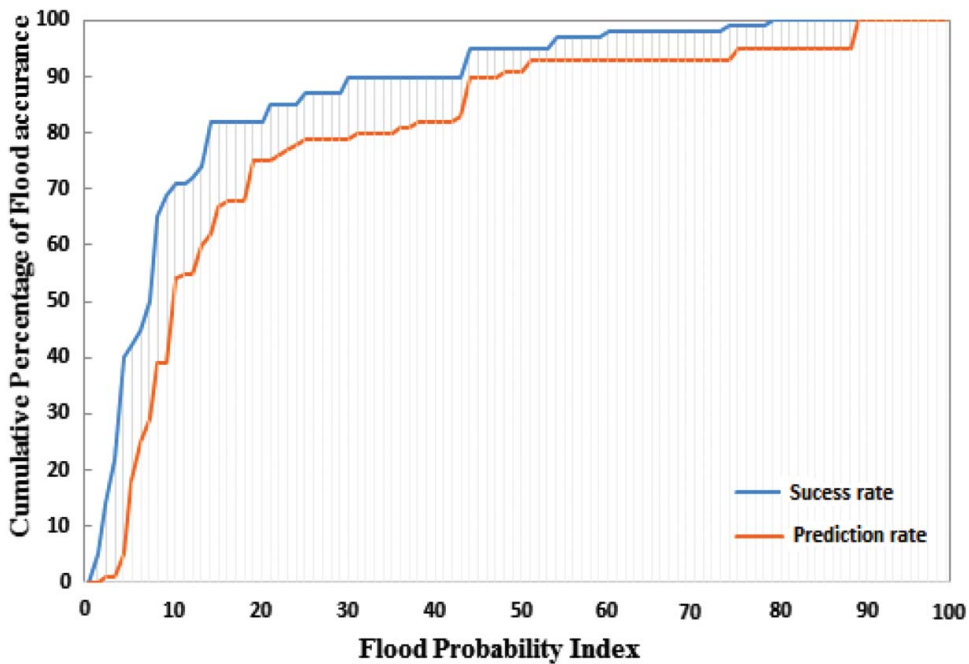


Figure 12. Area under the curve for success rate (89.7%) and prediction rate (78.9%).

NKVE, and the location of the study area at Sungai Damansara is highly susceptible to flood occurrence. The Sungai Damansara catchment was used in the case study, and 13 indices were selected to construct the evaluation system. In the case of topographical data, DEM was built from IFSAR images with a pixel size of $5 \text{ m} \times 5 \text{ m}$. Furthermore, LULC mapping, which was one of the effective parameters of flooding, and vulnerability assessment were extracted from WorldView-3 satellite imagery with 0.3 m spatial resolution. Each conditioning factor was optimized using the FR approach and entered as input for SVM modelling. The correlation between each flood-related factor and flood location showed that areas with a low elevation, mild slope, and flat curvature exhibit the highest probability of flood occurrence. The accuracy of the flood probability indices obtained from this ensemble method was validated using AUC. The estimated success rate and prediction rate of the applied method was at 89.7% and 78.9%, respectively. The most effective parameters that would trigger flood occurrence in the study area were rainfall and flood inundation depth. Thus, it was used as the triggering factor for flood hazard estimation. Rainfall data were derived from nine rainfall stations in and around the study area within the last 5 years and flood inundation depth was generated using hydraulic 2D HRS model. Furthermore, weights ranging from 1 to 5 were assigned to the most vulnerable elements, which were selected based on precise land-use information. As the main objective of this research, the risk level map was finally produced based on hazard and vulnerability indices. The reliability of the results obtained from this study was also verified in the field. Consequently, the ensemble method of FR and SVM can be efficiently used in flood hazard studies because of its simple structure and robust performance. Furthermore, the FR model is an excellent approach for ranking different classes of conditioning parameters. Therefore, each index map derived from this study can be helpful to planners and decision makers for flood management and planning in the study area.

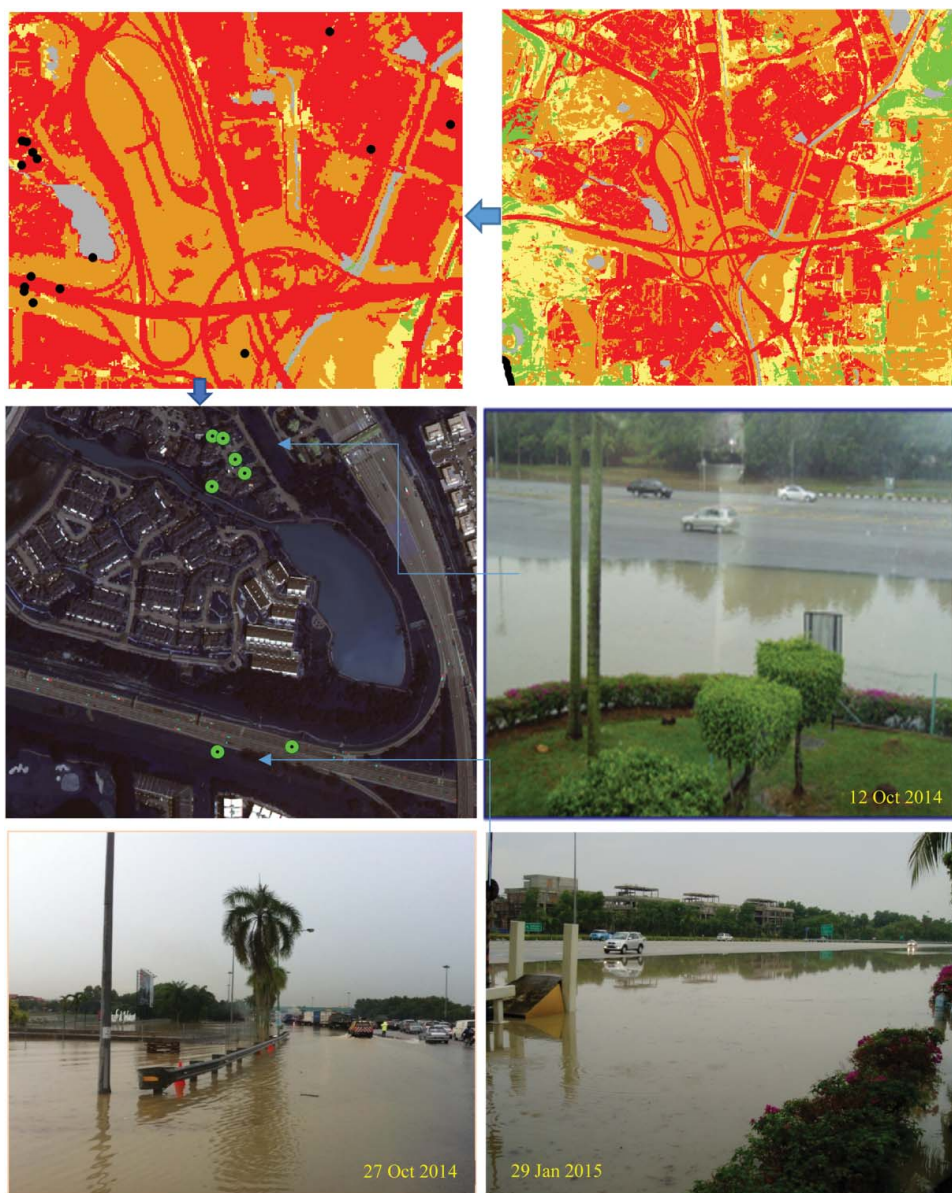



Figure 13. Field verification for Damansara river catchment.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Biswajeet Pradhan  <http://orcid.org/0000-0001-9863-2054>

References

- Akgun A, Dag S, Bulut F. 2007. Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models. *Environ Geol.* 54:1127–1143. doi:10.1007/s00254-007-0882-8.
- Althuwaynee OF, Pradhan B, Lee S. 2012. Application of an evidential belief function model in landslide susceptibility mapping. *Comput Geosci.* 44:120–135. doi:10.1016/j.cageo.2012.03.003.
- Apel H, Aronica GT, Kreibich H, Thieken AH. 2008. Flood risk analyses – how detailed do we need to be? *Nat Hazards.* 49:79–98. doi:10.1007/s11069-008-9277-8.
- Ayalew L, Yamagishi H. 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology.* 65:15–31. doi:10.1016/j.geomorph.2004.06.010.
- Botzen WJW, Aerts JCJH, van den Bergh JCJM. 2012. Individual preferences for reducing flood risk to near zero through elevation. *Mitig Adapt Strateg Glob Change.* 2:229–244. doi:10.1007/s11027-012-9359-5.
- Casulli, V. 2009. A high-resolution wetting and drying algorithm for free-surface hydrodynamics. *Int J Numerical Methods Fluids.* 60:391–408.
- Chau KW, Wu CL, Li YS. 2005. Comparison of several flood forecasting models in Yangtze River. *J Hydrol Eng.* 10: 485–491. doi:10.1061/(asce)1084-0699(2005)10:6(485).
- Choi J, Oh HJ, Won JS, Lee S. 2009. Validation of an artificial neural network model for landslide susceptibility mapping. *Environ Earth Sci.* 60:473–483. doi:10.1007/s12665-009-0188-0.
- Gokceoglu C, Sonmez H, Nefeslioglu HA, Duman TY, Can T. 2005. The 17 March 2005 Kuzulu landslide (Sivas, Turkey) and landslide-susceptibility map of its near vicinity. *Eng Geol.* 81:65–83. doi:10.1016/j.enggeo.2005.07.011.
- Horritt MS. 2006. A methodology for the validation of uncertain flood inundation models. *J Hydrol.* 326:153–165. doi:10.1016/j.jhydrol.2005.10.027.
- Rizeei HM, Saharkhiz MA, Pradhan B, Ahmad N. 2016. Soil erosion prediction based on land cover dynamics at the Semeniyih watershed in Malaysia using LTM and USLE models. *Geocarto Int.* 31:1158–1177. doi: 10.1080/10106049.2015.1120354.
- Huong HTL, Pathirana A. 2013. Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. *Hydrol Earth Syst Sci.* 17:379–394. doi:10.5194/hess-17-379-2013.
- Intergovernmental Panel on Climate Change, 2014. Summary for policy makers. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir MCTE, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea a.L.L.W.e.P.R. editors. *Climate change 2014: impact, adaptation, and vulnerability. Part a: global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change.* Cambridge (UK): Cambridge University Press; p. 1–32.
- Jebur MN, Pradhan B, Tehrany MS. 2014a. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sens Environ.* 152:150–165. doi:10.1016/j.rse.2014.05.013.
- Kar AK, Lohani AK, Goel NK, Roy GP. 2015. Rain gauge network design for flood forecasting using multi-criteria decision analysis and clustering techniques in lower Mahanadi river basin, India. *J Hydrol: Reg Stud.* 4:313–332. doi:10.1016/j.ejrh.2015.07.003.
- Kia MB, Pirasteh S, Pradhan B, Mahmud AR, Sulaiman WNA, Moradi A. 2011. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environ Earth Sci.* 67:251–264. doi:10.1007/s12665-011-1504-z.
- Lecca G, Petitdidier M, Hluchy L, Ivanovic M, Kussul N, Ray N, Thieron V. 2011. Grid computing technology for hydrological applications. *J Hydrol.* 403:186–199.
- Lee MJ, Kang J, Jeon S. 2012. Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. 2012 IEEE International Geoscience and Remote Sensing Symposium. doi:10.1109/igarss.2012.6351414.
- Manandhar B. 2010. Flood plain analysis and risk assessment of Lothar Khola, MSc Thesis, Phokara: Tribhuvan University; p. 64.
- Marjanović M, Kovačević M, Bajat B, Voženilek V. 2011. Landslide susceptibility assessment using SVM machine learning algorithm. *Eng Geol.* 123:225–234.
- Merz B, Kreibich H, Lall U. 2013. Multi-variate flood damage assessment: a tree-based data-mining approach. *Nat Hazard Earth Sys.* 13:53–64. doi:10.5194/nhess-13-53-2013.
- Meyer V, Scheuer S, Haase D. 2008. A multicriteria approach for flood risk mapping exemplified at the Mulde river, Germany. *Nat Hazards.* 48:17–39. doi:10.1007/s11069-008-9244-4.
- Moore ID, Wilson JP. 1992. Length-slope factors for the revised universal soil loss equation: simplified method of estimation. *J Soil Water Conserv.* 47:423–428.
- Müller A, Reiter J, Weiland U. 2011. Assessment of urban vulnerability towards floods using an indicator-based approach – a case study for Santiago de Chile. *Nat Hazards Earth Syst Sci.* 11:2107–2123. doi:10.5194/nhess-11-2107-2011.

- Neshat A, Pradhan B. 2015. An integrated DRASTIC model using frequency ratio and two new hybrid methods for groundwater vulnerability assessment. *Nat Hazards*. 76:543–563.
- Pradhan B. 2009. Groundwater potential zonation for basaltic watersheds using satellite remote sensing data and GIS techniques. *Open Geosci*. 1:120–129. doi:10.2478/v10085-009-0008-5.
- Pradhan B. 2010. Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. *J Spatial Hydrol*. 9:1–18.
- Pradhan B, Hagemann U, Tehrany, MS, Prechtel N. 2014. An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image. *Comput Geosci*. 63:34–43. doi:10.1016/j.cageo.2013.10.011.
- Samui P. 2008. Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Comput Geotech*. 35:419–427. doi:10.1016/j.compgeo.2007.06.014.
- Song S, Zhan Z, Long Z, Zhang J, Yao L. 2011. Comparative study of SVM methods combined with voxel selection for object category classification on fMRI Data. *PLoS ONE*. 6:e17191. doi:10.1371/journal.pone.0017191.
- Stefanidis S, Stathis D. 2013. Assessment of flood hazard based on natural and anthropogenic factors using analytic hierarchy process (AHP). *Nat Hazards*. 68:569–585. doi:10.1007/s11069-013-0639-5.
- Tehrany MS, Pradhan B, Jebur MN. 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J Hydrol*. 512:332–343. doi:10.1016/j.jhydrol.2014.03.008.
- Tehrany MS, Pradhan B, Jebur MN. 2013. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J Hydrol*. 504:69–79. doi:10.1016/j.jhydrol.2013.09.034.
- Tien Bui, D., Pradhan, B., Lofman, O., Revhaug, I., 2012a. Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naïve Bayes Models. *Math Probl Eng*. 2012:1–26. doi:10.1155/2012/974638.
- Tingsanchali T, Karim F. 2010. Flood-hazard assessment and risk-based zoning of a tropical flood plain: case study of the Yom River, Thailand. *Hydrol Sci J*. 55:145–161. doi:10.1080/02626660903545987.
- Tunusluoglu MC, Gokceoglu C, Nefeslioglu HA, Sonmez H. 2007. Extraction of potential debris source areas by logistic regression technique: a case study from Barla, Besparmak and Kapi mountains (NW Taurids, Turkey). *Environ Geol*. 54:9–22. doi:10.1007/s00254-007-0788-5.
- Vojinović Z, Abbott MB. 2012. Flood risk and social justice: from quantitative to qualitative flood risk assessment and mitigation. *Water Intell*. Online 11. London (UK): IWA Publishing. ISBN: 9781780400822. Horritt, 2006.
- Yalcin A, Reis S, Aydinoglu AC, Yomralioglu T. 2011. A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *CATENA*. 85:274–287. doi:10.1016/j.catena.2011.01.014.
- Yao X, Tham L, Dai F. 2008. Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. *Geomorphology*. 101:572–582.
- Zhou J, Deng W, Zou Q, Xiao J, Zhang Y, Hua W. 2013. Flood disaster evaluation model based on kernel dual optimization support vector machine. *Information Technol J*. 12:2412–2418. doi:10.3923/ijtj.2013.2412.2418.