



# Flood susceptibility analysis through remote sensing, GIS and frequency ratio model

Sailesh Samanta<sup>1</sup> · Dilip Kumar Pal<sup>1</sup> · Babita Palsamanta<sup>2</sup>

Received: 19 June 2017 / Accepted: 10 April 2018 / Published online: 21 April 2018  
© The Author(s) 2018

## Abstract

Papua New Guinea (PNG) is saddled with frequent natural disasters like earthquake, volcanic eruption, landslide, drought, flood etc. Flood, as a hydrological disaster to humankind's niche brings about a powerful and often sudden, pernicious change in the surface distribution of water on land, while the benevolence of flood manifests in restoring the health of the thalweg from excessive siltation by redistributing the fertile sediments on the riverine floodplains. In respect to social, economic and environmental perspective, flood is one of the most devastating disasters in PNG. This research was conducted to investigate the usefulness of remote sensing, geographic information system and the frequency ratio (FR) for flood susceptibility mapping. FR model was used to handle different independent variables via weighted-based bivariate probability values to generate a plausible flood susceptibility map. This study was conducted in the Markham riverine precinct under Morobe province in PNG. A historical flood inventory database of PNG resource information system (PNGRIS) was used to generate 143 flood locations based on "create fishnet" analysis. 100 (70%) flood sample locations were selected randomly for model building. Ten independent variables, namely land use/land cover, elevation, slope, topographic wetness index, surface runoff, landform, lithology, distance from the main river, soil texture and soil drainage were used into the FR model for flood vulnerability analysis. Finally, the database was developed for areas vulnerable to flood. The result demonstrated a span of FR values ranging from 2.66 (least flood prone) to 19.02 (most flood prone) for the study area. The developed database was reclassified into five (5) flood vulnerability zones segmenting on the FR values, namely very low (less than 5.0), low (5.0–7.5), moderate (7.5–10.0), high (10.0–12.5) and very high susceptibility (more than 12.5). The result indicated that about 19.4% land area as 'very high' and 35.8% as 'high' flood vulnerable class. The FR model output was validated with remaining 43 (30%) flood points, where 42 points were marked as correct predictions which evinced an accuracy of 97.7% in prediction. A total of 137292 people are living in those vulnerable zones. The flood susceptibility analysis using this model will be very useful and also an efficient tool to the local government administrators, researchers and planners for devising flood mitigation plans.

**Keywords** Remote sensing · GIS · Flood vulnerability analysis · Frequency ratio · Markham river

## Introduction

High intensity downpours in a region often lead to flooding in the downstream areas. Floods happen when overland flow of waters inundates land (Merz et al. 2010). Natural disasters, like floods, are causing massive damages to natural and human resources (Du et al. 2013; Youssef et al. 2011). An average of 140 million people is affected per year due to flooding (WHO 2003). In respect of socioeconomic and environmental consequences, widespread flood analysis is very significant (Markantonis et al. 2013). Control of a flood and prevention measures are necessary to reduce potential damages to natural resources, agriculture, infrastructure etc. (Billa et al. 2006; Huang et al. 2008). Therefore, analysis of

✉ Sailesh Samanta  
rsgis.sailesh@gmail.com

Dilip Kumar Pal  
dkpal200090@gmail.com

Babita Palsamanta  
mundey2@gmail.com

<sup>1</sup> Department of Surveying and Land Studies, The PNG University of Technology, Private Mail Bag, Lae, Morobe, Papua New Guinea

<sup>2</sup> PNG University of Technology Campus, Private Mail Bag, Lae, Morobe, Papua New Guinea

flood susceptibility is an important task for early warning system, emergency services towards management strategies of prevention and mitigation of the future flood episodes (Tehrany et al. 2015). There are several comprehensive tools available and used by many research organizations worldwide, for example HAZUS, a GIS-based natural hazard analysis tool developed for assessing flood hazard; HEC-FDA, a computer program to assist crop engineers through vulnerability analysis of flood risk reduction policy. RS and GIS techniques offer a suitable platform to manipulate and analyse all relevant information in order to easily demarcate suitable hazard zones (Khan et al. 2008; Saha et al. 2005; Wang et al. 2013; Pourghasemi et al. 2014). RS and GIS techniques are useful widely to demarcate and assess flood-related damages (Pradhan et al. 2014; Patel and Srivastava 2013) caused by excessive rain in a catchment area or sea wave surges in the coastal regions. During last decade, several models were applied for estimation of flood by many researcher and scientists (Kisi et al. 2012; Talei et al. 2010). Alternatively, the conventional flood modelling methods were not reliable for accurate prediction (Tehrany et al. 2014b; Li et al. 2012). Currently, geospatial techniques provide a wide range of data sources for the modelling of flood (Wanders et al. 2014). Analytical hierarchy process (AHP) is one of the most popular and satisfactory methods in disaster modelling like flood monitoring, mapping and analysing complex problems (Billa et al. 2006; Yalcin 2008; Chen et al. 2011). Apart from AHP, multi-criteria decision support approach (MCDA) (Samanta et al. 2016a), weights of evidence (WofE) (Rahmati et al. 2016a), logistic regression (LR) (Tehrany et al. 2014a, b, 2015), adaptive neuro-fuzzy interface system (Sezer et al. 2011), artificial neural networks (ANN) (Pradhan and Buchroithner 2010; Tiwari and Chatterjee 2010) and FR model (Lee et al. 2012; Rahmati et al. 2016b; Liao and Carin 2009) were the other known acceptable models for hazard analysis. ANN method was used in prediction of flood, where researchers attempted to show relationship between conditioning parameters and the outcome (Pradhan and Buchroithner 2010). It was reported that ANN method can handle all inputs which are uncertain to extract meaningful information (Lohani et al. 2012). MCDA, RS and GIS techniques are extremely useful in reliable and accurate analysis and mapping of plausible flood prone zones. MCDA approach is suitable for flood analysis and mapping in the no-data regions and could be practical for local planners in mitigation of flood. AHP model was applied in China for flood diversion (Zou et al. 2013). According to Chen et al. (2011), the weakness of AHP model was correlated to its enslavement on the information provided by experts, which is the main source of uncertainty. FR model may be considered as an important method which is easily understandable and can be used to produce acceptable flood risk analysis and mapping (Liao and Carin 2009).

WofE and FR model are relatively new for flood vulnerability modelling, and they were widely used in other natural hazards like landslide susceptibility mapping (Rahmati et al. 2016b). Both models are almost similar and yield comparable results for flood susceptibility mapping. Flood risk map through FR model was expected to be used in programmes to reduce the flood and its damage (Lee et al. 2012).

Modelling of flood is a complex exercise where a lot of factors are supposed to be considered. RS technique provides a significant contribution in flood mapping and risk assessment. RS and GIS are quick and more efficient, which can provide the best opportunity to capture, store, combine, manipulate, retrieve, analyse and display the information for the determination of potential hazard areas. This research is an ensemble method, which proves the efficiency in GIS-based flood modelling. To estimate flood probabilities, the frequency ratio (FR) approach was combined with RS and GIS.

The main goal of the present research is to examine the usefulness of Remote Sensing (RS), GIS and the frequency ratio (FR) models for flood susceptibility analysis and mapping in the Markham river basin under Morobe province, Papua New Guinea. The main aim of this study is to identify and map out flood risk zones in the Markham river basin. The objectives of this research are to create wall-to-wall data sets that are considered as input into the FR model to categorize potential flood prone area, create a flood hazard map of the Markham river basin and also to carry out impact analysis which can be useful to local government administrators, researchers and planners for devising flood mitigation plans.

## Study location and materials used

The study was carried out in the final basin (basin-14) of Markham, which is located in the Morobe province of PNG and encloses an area of 1806.85 km<sup>2</sup>. The longitudinal and latitudinal extensions of the study area are 146.09–147.04° east and 6.23–6.78° south, respectively (Fig. 1). The climate is tropical humid with about 4200 mm of average annual rainfall in the study area. Markham river emanates from the runoff contributed by 12,450 km<sup>2</sup> catchment (Fig. 1a) with huge bed load (Tilley et al. 2006). This is the fourth longest river in PNG born at Finisterre range (approximately 457 m altitude) and flows into the Huon Gulf near to the downtown of Lae, which is the second largest city of PNG after 180 km of checkered path (Samanta et al. 2016b). The upper basin is dominated by natural forests, steep slopes and rugged terrain (Pal et al. 2012; Solin 2012). Mining activities for alluvial gold extraction on the river path and logging activities in the surroundings areas result in accelerated soil erosion within the basin area. Final 125 km downstream flow of the

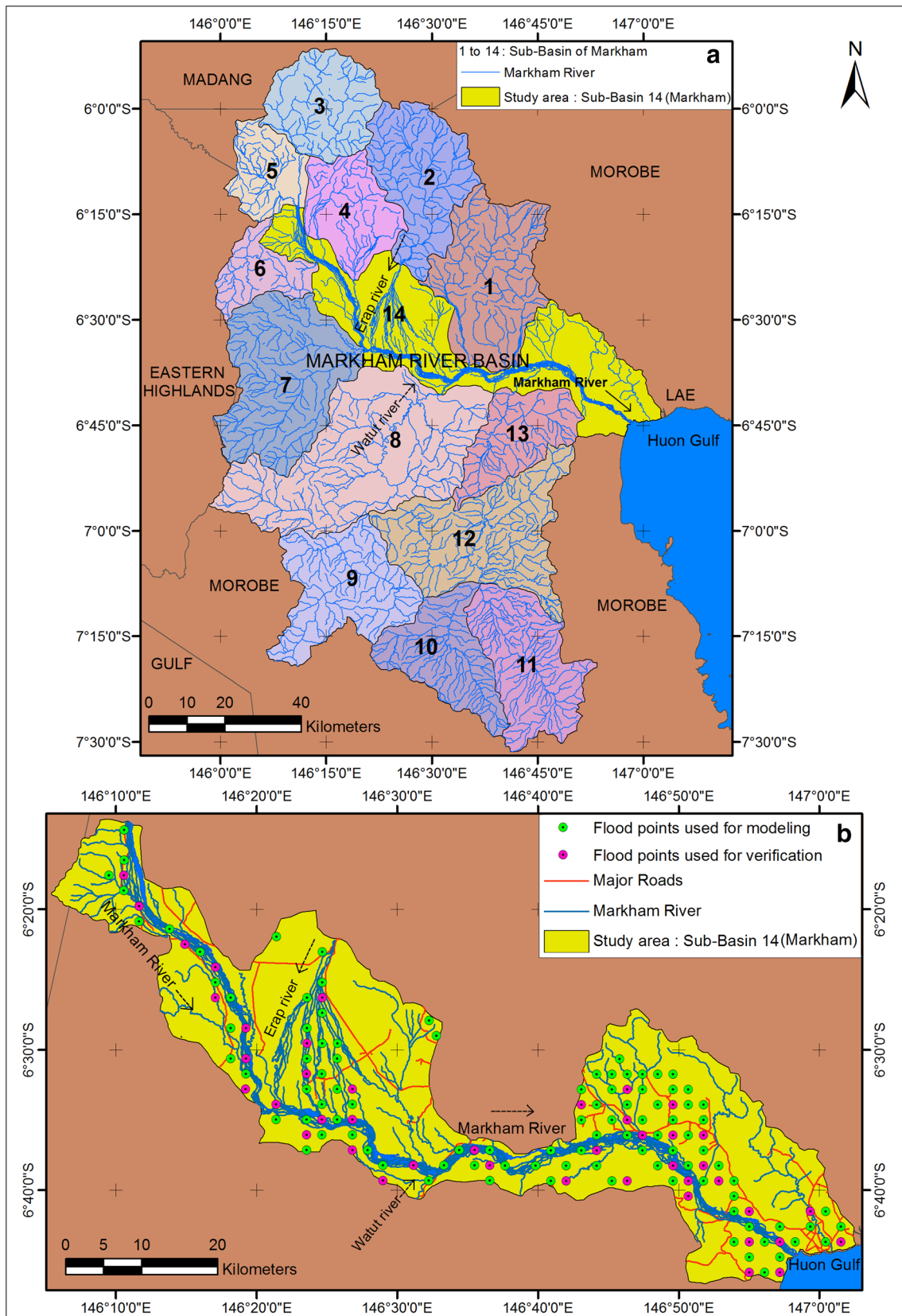


Fig. 1 Location map of the study area: (a) Entire Markham river basin and (b) The study area-sub-basin 14

Markham river which is covered by sub-basin number 14 was selected for this study (Fig. 1b).

The Markham river catchment in Morobe province of PNG is exposed to flood due to intensive rainfall, peak discharge and physiographic conditions (Tilley et al. 2006; Samanta et al. 2016a). During 2012, higher magnitude of floods was measured by the global flood detection system (Kugler et al. 2007). On March of 2004, Markham riverine zones experienced an overwhelming flood on the spate of Markham river. Earlier on, during the month of February (2004), 120-m-long road on Lae-Bulolo road was washed out just 1.5 km away towards upstream from Markham bridge. The estimated peak discharge was about 2600 m<sup>3</sup>/s to 3200 m<sup>3</sup>/s with an average velocity as 3.4 m/s (Tilley et al. 2006).

To perform flood susceptibility analysis and risk assessments, researcher suggested different factors which are not fixed (Tehrany et al. 2015). There are some common conditioning factors which indicate their role in flood mapping. Recent studies by Rahmati et al. (2016a, b) have achieved high accurate results, where they used the least number of independent parameters. Total of 15 parameters are examined as a preliminary analysis. Demarcation of flood susceptibility zones was carried out using ten (10)-folds of geospatial data sets, viz. land use and land cover (LULC), elevation, slope, topographic wetness index (TWI), surface runoff, landform, lithology, distance from the main river, soil texture and soil drainage. These geospatial layers are selected after consulting local hydrological and natural disaster experts based on their effectiveness in creating a flood. Wall-to-wall geospatial database was developed from remotely sensed data sets like satellite image and digital elevation model and National GIS database of PNG. All data sets were rectified carefully using Erdas imagine software. UTM projection, zone 55S and WGS-84 datum were selected for the image and map registration purpose. Advanced space-borne thermal emission and reflection radiometer (ASTER) provides digital elevation model (DEM) with the spatial resolution of 30 m that was used in this research. Elevation and slope database were derived from the ASTER DEM. LULC was derived from optical bands with false colour combination

of Landsat 8, operational land imager (OLI). A flood inventory database was prepared using PNGRIS national database, field investigation and by examining remote sensing data. Every detail of the data sets used in FR model is given in Table 1.

## Methodology

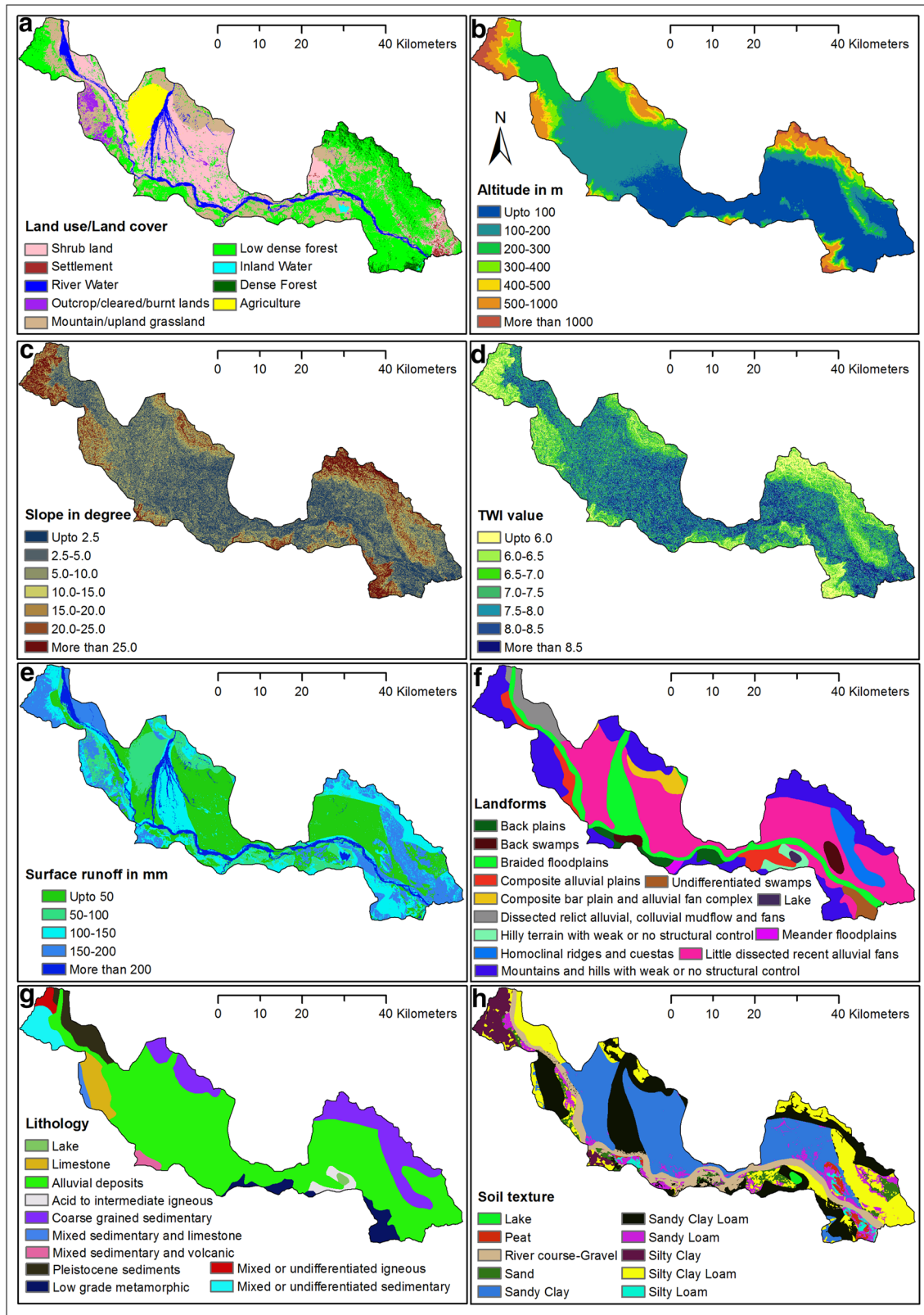
It is essential to analyse the occurrence of historical flood events to estimate future flood (Manandhar 2010). So the flood inventory database is the essential factor for flood susceptibility mapping. Flood inventory map was prepared after generating 143 flooded points through “create fishnet” analysis of PNGRIS national database, field investigation and satellite data before and after flood events. Seventy percent (70%) of total, i.e. 100 flood points, were selected randomly as the training data set for flood modelling and the rest 30% or 43 points were used for validating the flood model (Rahmati et al. 2016b) (Fig. 1b).

Selection of effectual parameters is vital to produce a flood hazard map in any catchment (Kia et al. 2012). It is always tricky to choose factors unanimously for use in flood susceptibility mapping (Tehrany et al. 2014a, b). Flood-related geospatial database like land use/land cover, elevation, slope, topographic wetness index (TWI), surface runoff, landform, lithology, distance from the main river, soil texture and soil drainage were prepared using ArcGIS and Erdas Imagine software.

LULC directly or indirectly influence infiltration, evapotranspiration and surface runoff generation. LULC map was prepared (Fig. 2a) from the Landsat-8 OLI satellite imagery through supervised classification technique (Samanta et al. 2011). The altitude and slope are important parameters in flood risk and vulnerability mapping. Variations of elevations have a definitive impact on climate characteristics (Samanta et al. 2012). Such different rainfall and temperature regimes engendered varied vegetation and soil forms (Aniya 1985). Slope controls the surface runoff, the ferocity of water flow aggravating soil erosion (Adiat et al. 2012) as well as vertical percolation (Youssef

**Table 1** Data sets used for flood vulnerability mapping

Sl. No.	Data description	Data type/resolution	Year	Source
1	LANDSAT-8, OLI	15 m Pan-Sharp	2013	Department of Surveying and Land Studies, The PNG University of Technology
2	ASTER-DEM	30 m	2001	
3	Rainfall,	Polygon/shape file	2009	PNGRIS—PNG Resource Information System
4	Soil texture and drainage			
5	Lithology			
6	Landforms			
7	Historical flood inventory database	Point location		



**Fig. 2** Parameters used for FR modelling: **(a)** land use/land cover, **(b)** altitude, **(c)** slope, **(d)** TWI index, **(e)** surface runoff, **(f)** landforms, **(g)** lithology and **(h)** soil texture characteristics of the study area

et al. 2011). Both altitude and slope map were prepared using ASTER data in ArcGIS 3-D analysis algorithms (Fig. 2b, c).

Topographic wetness index (TWI) refers to spatial distribution of wetness and controls the overland flow of water. TWI has significant impact on flood mapping. TWI was calculated based on Eq. 1 (Beven and Kirkby 1979; Regmi et al. 2010; Qin et al. 2011), and spatial distribution map of TWI was prepared using ArcGIS (Fig. 2d).

$$TWI = \text{Ln}\left(\frac{a}{\tan B}\right), \quad (1)$$

where  $a$  is the specific catchment area [ $a = A/L$ , total basin area ( $A$ ) divided by length of contour ( $L$ )], and  $B$  is referred to the slope in degree.

Overland flow of water, called surface runoff, occurs in full throttle in the aftermath of saturated infiltration when the surplus water minus saturated infiltration component, emanating from the storm, melt water, etc. keep flowing over the Earth's surface (Pal and Samanta 2011). Surface runoff in urban areas is normally reinforced owing to lack of infiltration surface, which gives rise to pernicious urban flooding. Surface runoff due to storm rainfall is very significant for forecasting floods which are very sudden, flashy and of short duration (Pal and Samanta 2011). Surface runoff database was generated (Fig. 2e) based on soil conservation service (SCS) model (Eqs. 2–4).

$$Q = (P - Ia)^2 / (P - Ia + S), \quad (2)$$

where  $Q$  is actual surface runoff in mm;  $P$  is storm rainfall (mm);  $S$  is the potential maximum retention (mm), and  $Ia$  is  $0.4S$  [Initial abstraction (mm)].

So the modified form of Eq. 2 can be expressed in Eq. 3.

$$Q = (P - 0.4S)^2 / (P + 0.6S). \quad (3)$$

To calculate the value of potential maximum retention ( $S$ ) of Eq. 3, another simple Eq. 4 was used.

$$S = (25400 / CN) - 254, \quad (4)$$

where  $CN$  is curve number of hydrologic soil cover complex, which happens to be a function of soil type, land cover and antecedent moisture condition (SCS 1972; Kumar et al. 1991; Rao et al. 1996).

Flood generally occurs near to the bank of the river and inundates low-lying flood plain areas. Distance from the river has significant impact on the flood and its magnitude. Distance from main river was considered as one parameter which was developed through proximity analysis in ArcGIS. Infiltration varies based on the spatial distribution of soil texture and it controls overland flows and inundation. Database on landform, lithology, soil texture (Fig. 2f–h)

and soil drainage was prepared based on PNRIS data sets after verifying them based on field observations.

It is essential to analyse past flood record to estimate the future flood event in any area (Manandhar 2010). Therefore, mapping the flood locations from past episodes in the study area is instrumental in elucidating the correlation among the flooding and the condition factors. It is essential to prepare an inventory database with high precision (Jebur et al. 2013). FR model was adopted for this study chosen from a plethora of bivariate statistical techniques. The approach offers a quantitative relationship between the 'frequency of flood episodes' and various conditioning parameters. The FR index was calculated using the following Eq. 5 (Tehrany et al. 2014a, b).

$$FSI = \sum FR, \quad (5)$$

where  $FSI$  is the flood susceptibility index and  $FR$  is the frequency ratio for each factor.

The  $FR$  can be expressed as the proportion of flooded area in the total study area (Eq. 6); it is the ratio of the probabilities of the 'flooded' to 'not flooded' area for a given attribute (Bonham-Carter 1994).

$$FR = (E / F) / (M / L), \quad (6)$$

where  $E$  is the number of flood episodes for each factor;  $F$  is the total number of flood episodes;  $M$  is the histogram of a class;  $L$  is the total histogram of the study area.

The frequency ratio model was used to establish the correlation between historical flood locations and the probable supporting factors.  $FR$  value indicates (Table 2) the types of correlation between factors and floods. A  $FR$  value lower than 1 indicates weak correlation; on the other hand a value of more than 1 refers to strong correlations. The overall methodology that was applied for flood susceptibility mapping is shown as a flowchart in Fig. 3.

## Results and discussion

There are many independent variables, called conditioning factors that play specific role in order to perform flood susceptibility mapping (Pradhan 2010; Pourghasemi et al. 2012; Kia et al. 2012). Spatial distribution and statistical database for all ten (10) conditioning factors namely, LULC, elevation, slope, topographic wetness index (TWI), surface runoff, landform, lithology, distance from the main river, soil texture and soil drainage were constructed with their subclasses (Fig. 2 and Table 2). Classification of satellite data was done (Samanta et al. 2012) considering nine (9) land use/land cover categories, namely dense forest, low dense forest, shrub land, outcrop/cleared/burnt lands, mountain/upland grassland, settlement, inland water, river water and

**Table 2** Conditioning parameters used for flood vulnerability mapping through FR model

Value	Class name or description	Histogram	% of Histogram	Flood numbers	% flood numbers	Frequency Ratio
<b>Land use/land cover</b>						
1	Dense forest	50700	2.5	1	1.0	0.40
2	Low dense forest	703420	35.0	33	33.0	0.94
3	Shrub land	533068	26.6	26	26.0	0.98
4	Outcrop/cleared/burnt lands	37806	1.9	0	0.0	0.00
5	Mountain/upland grassland	433316	21.6	28	28.0	1.30
6	Settlement	18728	0.9	0	0.0	0.00
7	Inland water	5406	0.3	0	0.0	0.00
8	River water	133552	6.7	11	11.0	1.65
9	Agriculture	91612	4.6	1	1.0	0.22
<b>Elevation (in m)</b>						
1	Up to 100 m	746583	37.2	59	59.0	1.59
2	100–200 m	544304	27.1	25	25.0	0.92
3	200–300 m	301652	15.0	11	11.0	0.73
4	300–400 m	120706	6.0	5	5.0	0.83
5	400–500 m	69673	3.5	0	0.0	0.00
6	500–1000 m	177557	8.8	0	0.0	0.00
7	More than 1000 m	47133	2.3	0	0.0	0.00
<b>Slope (in degree)</b>						
1	up to 2.5°	336363	16.8	21	21.0	1.25
2	2.5–5.0°	513214	25.6	40	40.0	1.56
3	5.0–10.0°	563975	28.1	29	29.0	1.03
4	10.0–15.0°	209103	10.4	2	2.0	0.19
5	15.0–20.0°	131856	6.6	5	5.0	0.76
6	20.0–25.0°	100073	5.0	2	2.0	0.40
7	More than 25°	153024	7.6	1	1.0	0.13
<b>Topographic wetness index (TWI)</b>						
1	up to 6.0	129577	6.5	1	1.0	0.15
2	6.0–6.5	226488	11.3	6	6.0	0.53
3	6.5–7.0	250979	12.5	3	3.0	0.24
4	7.0–7.5	387731	19.3	16	16.0	0.83
5	7.5–8.0	449732	22.4	32	32.0	1.43
6	8.0–8.5	299489	14.9	28	28.0	1.88
7	More than 8.5	263612	13.1	14	14.0	1.07
<b>Surface runoff (in mm)</b>						
1	Up to 50 mm	682498	34.0	41	41.0	1.21
2	50–100 mm	281239	14.0	8	8.0	0.57
3	100–150 mm	486606	24.2	28	28.0	1.16
4	150–200 mm	406985	20.3	13	13.0	0.64
5	More than 200 mm	150280	7.5	10	10.0	1.34
<b>Landform</b>						
1	Dissected relict alluvial, colluvial mudflow and fans	75020	3.7	0	0.0	0.00
2	Mountains and hills with weak or no structural control	476893	23.8	0	0.0	0.00
3	Braided floodplains	296938	14.8	36	36.0	2.43
4	Composite alluvial plains	94495	4.7	10	10.0	2.12
5	Composite bar plain and alluvial fan complex	46996	2.3	1	1.0	0.43
6	Little dissected recent alluvial fans	779505	38.8	41	41.0	1.06
7	Homoclinal ridges and cuestas	77006	3.8	0	0.0	0.00
8	Back plains	46935	2.3	4	4.0	1.71

**Table 2** (continued)

Value	Class name or description	Histogram	% of Histogram	Flood numbers	% flood numbers	Frequency Ratio
9	Back swamps	40223	2.0	5	5.0	2.50
10	Hilly terrain with weak or no structural control	27365	1.4	0	0.0	0.00
11	Lake	6416	0.3	0	0.0	0.00
12	Meander floodplains	8914	0.4	1	1.0	2.25
13	Undifferentiated swamps	30902	1.5	2	2.0	1.30
<b>Lithology</b>						
1	Pleistocene sediments	75020	3.7	0	0.0	0.00
2	Coarse grained sedimentary	290647	14.5	0	0.0	0.00
3	Alluvial deposits	1344906	67.0	100	100.0	1.49
4	Mixed or undifferentiated igneous	19162	1.0	0	0.0	0.00
5	Mixed or undifferentiated sedimentary	61629	3.1	0	0.0	0.00
6	Limestone	74653	3.7	0	0.0	0.00
7	Mixed sedimentary and limestone	14059	0.7	0	0.0	0.00
8	Mixed sedimentary and volcanic	16871	0.8	0	0.0	0.00
9	Acid to intermediate igneous	27385	1.4	0	0.0	0.00
10	Lake	6416	0.3	0	0.0	0.00
11	Low grade metamorphic	76860	3.8	0	0.0	0.00
<b>Distance from river in m</b>						
1	up to 1000 m	393500	19.6	26	26.0	1.33
2	1000–2000 m	243382	12.1	19	19.0	1.57
3	2000–3000 m	206855	10.3	9	9.0	0.87
4	3000–4000 m	175027	8.7	13	13.0	1.49
5	4000–5000 m	149073	7.4	4	4.0	0.54
6	5000–6000 m	133660	6.7	4	4.0	0.60
7	More than 6000 m	706111	35.2	25	25.0	0.71
<b>Soil texture</b>						
1	Sandy clay loam	439768	21.9	20	20.0	0.91
2	Sandy loam	129127	6.4	13	13.0	2.02
3	Silty clay loam	377624	18.8	12	12.0	0.64
4	Peat	22672	1.1	1	1.0	0.89
5	Silty loam	22699	1.1	4	4.0	3.54
6	Sandy clay	634345	31.6	28	28.0	0.89
7	River course-gravel	216519	10.8	19	19.0	1.76
8	Sand	66418	3.3	2	2.0	0.60
9	Lake	6567	0.3	0	0.0	0.00
10	Silty clay	91869	4.6	1	1.0	0.22
<b>Soil drainage</b>						
1	Well drained	1631997	81.3	62	62.0	0.76
2	Waterlogged (swampy)	45371	2.3	5	5.0	2.21
3	Poorly to very poorly drained	105842	5.3	14	14.0	2.66
4	River course	216519	10.8	19	19.0	1.76
5	Imperfectly drained	1312	0.1	0	0.0	0.00
6	Lake	6567	0.3	0	0.0	0.00

agriculture land (Fig. 2a). The predominant land cover in the study area is 'low density forest' (35%) in the eastern region and some pockets of north-west and southern region of the study area. The 'river class' along the centre (main river) and north (Erap river) of the study area covers only 6.7% of the

study area, but the FR value was calculated as 1.65 which is highly correlated to flood (Table 2). The elevation ranged from 0 m to 1790 m of the study area. The spatial distribution of the elevation map was prepared after reclassifying the elevation layer into seven (7) classes (< 100, 100–200,



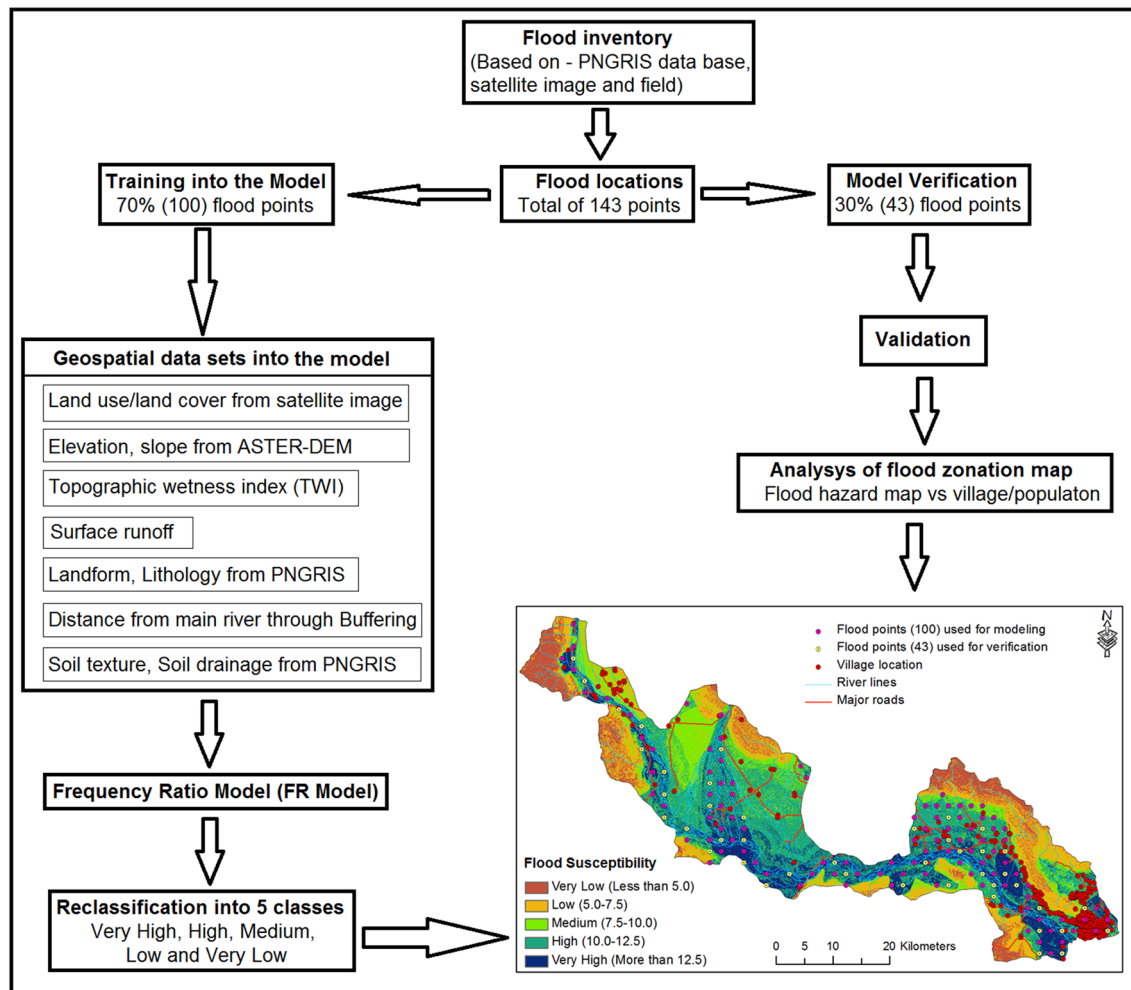
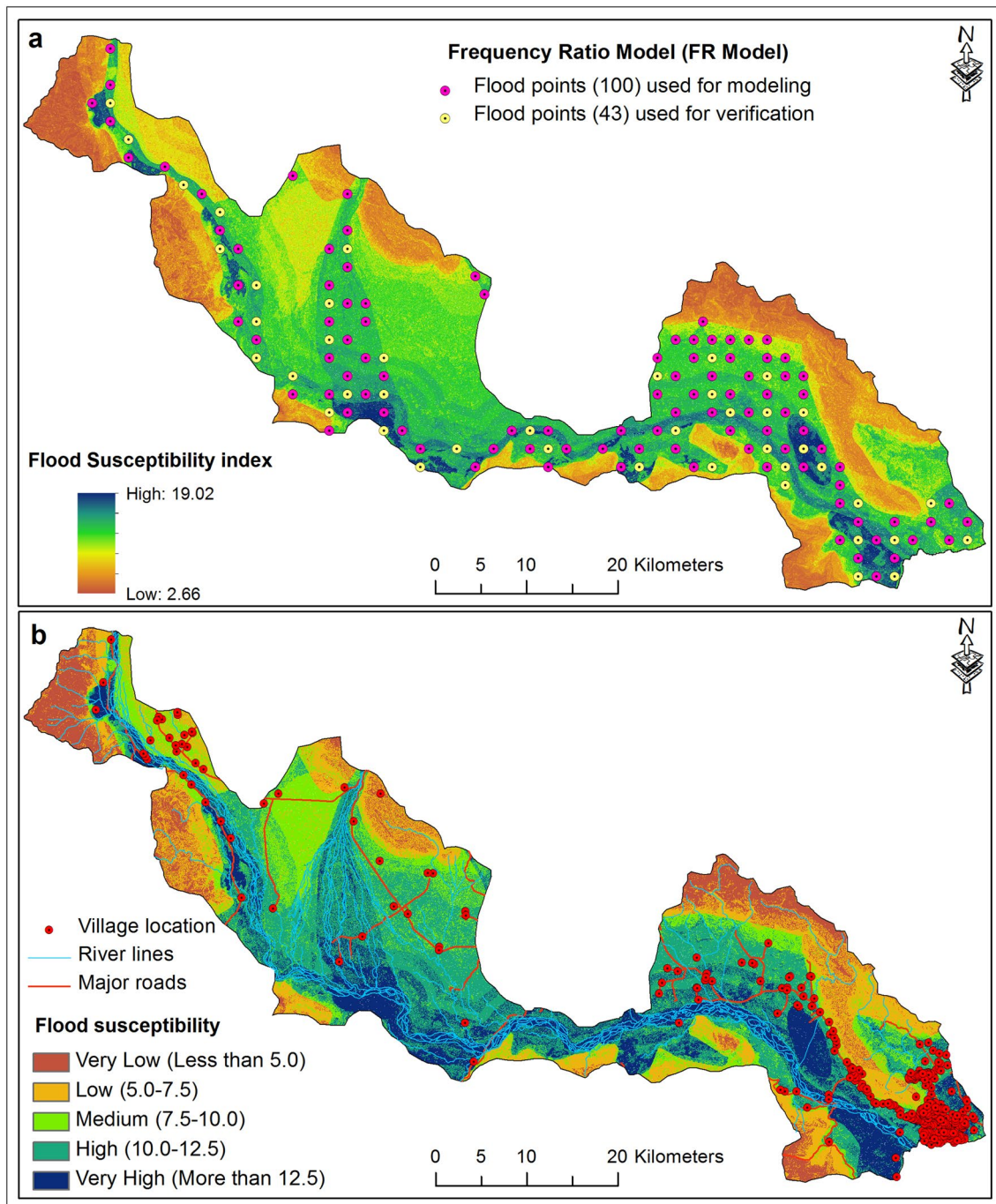


Fig. 3 Methodological flow chart of flood susceptibility mapping

200–300, 300–400, 400–500, 500–1000 and > 1000 m) where brown colour indicates maximum altitude found in north, north-west, and blue indicated < 100 m altitude zone in the east and south-east part in the study area (Fig. 2b). Up to 100 m altitude zone is the predominant class covering 37.2% of the total area, which refers to a higher FR value (1.59) than other classes (Table 2). Slope (in degree) was calculated from DEM data and reclassified into seven (7) classes, namely < 2.5°, 2.5–5°, 5–10°, 10–15°, 15–20°, 20–25° and more than 25°. The flat or lower slope gradient (< 5°) area is situated on both sides of the river as shown with blue colour (Fig. 2c). Lower the slope gradient, the more is the possibility of flooding and flood events (Rahmati et al. 2016a). Two lower slope gradient classes, namely < 2.5° and 2.5–5°, indicates higher FR value of 1.25 and 1.56, respectively, whereas > 25° slope area indicated lower FR value of 0.13 (Table 2).

Spatial database on TWI was calculated and categorized into seven (7) classes (< 6.0, 6.0–6.5, 6.5–7.0, 7.0–7.5,

7.5–8.0, 8.0–8.5 and > 8.5) as shown in Fig. 2d. The highest TWI was recorded in the middle part of the study area represented with blue colour. Higher TWI value refers higher chances of flooding in a watershed (Rahmati et al. 2016b). The FR value was calculated as 1.88 where TWI varied from 8.0 to 8.5, and 0.15 for the class with TWI < 6.0, respectively (Table 2). Surface runoff map was generated based on storm rainfall of 229.6 mm during 21–23 October 2012. Maximum surface runoff was calculated as 229 mm along the river class. High surface runoff during a storm indicates high probability of flood (Pal and Samanta 2011). The surface runoff database was reclassified into five (5) classes, namely < 50, 50–100, 100–150, 150–200, and > 200 mm (Fig. 2e). FR value was calculated as 1.34 for highest runoff category (> 200 mm) (Table 2). Entire study area has been categorized into thirteen (13) landform units, namely (i) dissected relict alluvial, colluvial mudflow and fans, (ii) mountains and hills with weak or no structural control, (iii) braided floodplains, (iv) composite alluvial plains,



**Fig. 4** Output map through FR model: (a) flood susceptibility index with all flood points and (b) overlay of villages and road on flood susceptibility zones

(v) composite bar plain and alluvial fan complex, (vi) little dissected recent alluvial fans, (vii) homoclinal ridges and cuestas, (viii) back plains, (ix) back swamps, (x) hilly terrain with weak or no structural control, (xi) lake, (xii) meander floodplains and (xiii) undifferentiated swamps (Fig. 2f). The back swamps and meander floodplains have higher FR value of 2.50 and 2.43, respectively (Table 2).

The lithology of the study area has been categorized into (i) Pleistocene sediments, (ii) coarse grained sedimentary, (iii) alluvial deposits, (iv) mixed or undifferentiated igneous, (v) mixed or undifferentiated sedimentary, (vi) limestone, (vii) mixed sedimentary and limestone, (viii) mixed sedimentary and volcanic, (ix) acid to intermediate igneous, (x) lake and (xi) low grade metamorphic (Fig. 2g). Lithology is

**Table 3** Spatial distribution (statistics) of flood vulnerability classes in the study area

Sl. no.	Flood vulnerable class	FR value range	Histogram	% area
1	Very low	<5.0	206164	10.3
2	Low	5.0–7.5	353062	17.6
3	Medium	7.5–10.0	340554	17.0
4	High	10.0–12.5	718863	35.8
5	Very high	> 12.5	388965	19.4

an important conditioning parameter in flooding because it has a direct influence on land permeability and thus surface runoff (Haghizadeh et al. 2017). The maximum FR value of 1.49 was recorded in alluvial deposits which covers 67% of the total study area (Table 2). Flood intensity became less in those locations far away from the river and the risk was higher in areas near to the river bank. Distances from the river in the range from 4000 to 5000, 5000 to 6000 and > 6000 m have a low probability of flooding, whereas distances in the range of < 1000 and 1000–2000 m together indicate highest FR values (1.42), which demonstrates the highest flood event (Haghizadeh et al. 2017). Soil texture and soil drainage are very important factors in flood susceptibility mapping. Well-drained soils produce less surface runoff than poorly drained soil group (Pal et al. 2012). The soil drainage database was generated with six (6) categories, namely (i) well drained, (ii) waterlogged (swampy), (iii) poorly to very poorly drained, (iv) river course, (v)

imperfectly drained and (vi) lake. Within the soil drainage factor, poorly to very poorly drained class had the highest FR value of 2.66, followed by waterlogged (swampy) area with the FR value of 2.21. Finally, the soil texture map was generated for the study area where ten (10) soil texture classes could be found (Samanta et al. 2016b), namely sandy clay loam, sandy loam, silty clay loam, peat, silty loam, sandy clay, river course-gravel, sand, lake and silty clay (Fig. 2h). The FR value was calculated as 3.54 for sandy loam and 0.22 for silty clay (Table 2).

Thus, the rating of each subclass of all conditioning parameter was generated based on the FR values as shown in Table 2. FR value varied from 0 to 3.54 in the study area. Calculated FR values were indicated as weak (<1) to strong (> 1) correlations with flood occurrence (Lee et al. 2012). Finally, based on FR model in Eq. 5, the flood probability database was developed (Fig. 4a). FR value in the model output varied from 2.66 to 19.02. Greater FR value indicated the higher probability to flood occurrences. The developed database was reclassified into five (5) different flood susceptibility zones, namely very low (less than 5.0), low (5.0–7.5), moderate (7.5–10.0), high (10.0–12.5) and very high susceptibility (more than 12.5) (Fig. 4b). The result indicates that about 19.4% land area are demarcated as very high, 35.8% as high, 17.0% as moderate, 17.6% as low and 10.3% as very low flood vulnerable class (Table 3). High to very high vulnerable classes are mostly located along the middle part of the study area (Fig. 4b). These high to very high flood susceptibility zones are characterized with higher runoff potentiality, poorly to very poorly drained soil, alluvial deposits, braided flood plain, lower slope

**Table 4** Calculation of prediction accuracy and success rate for the flood susceptibility analysis

Sl. no.	Susceptibility class	Verification (30% flood points)	Accurate (high to very high class)	Prediction accuracy	Training (70% flood points)	Success (high to very high class)	Success rate
1	Very low	0	42	0.977	0	94	0.94
2	Low	0		or	0		or
3	Medium	1		97.7%	6		94.0%
4	High	20			45		
5	Very high	22			49		
<b>Total</b>		<b>43</b>			<b>100</b>		

**Table 5** Flood susceptibility zone and the risk factor

	Susceptibility class	FR value range	Village	% of village under flood vulnerability	Total population
1	Very low	< 5.0	1	0.3	733
2	Low	5.0–7.5	16	4.5	5021
3	Medium	7.5–10.0	83	23.2	30466
4	High	10.0–12.5	183	51.1	93860
5	Very high	> 12.5	75	20.9	43432

gradient, lower elevation and closer to the main river, which are the important conditioning factors for flood susceptibility mapping using the FR model.

There are many models as proposed by different researchers, but it is very important to evaluate the accuracy and success rate to validate the model used for flood susceptibility analysis (Chung and Fabbri 2003; Tien Bui et al. 2012). The modelled output through FR model is validated in regard of success rate and prediction accuracy. The value of 1.0 represented the highest accuracy, which indicates the capability of the model in predicting natural hazards without any bias (Pradhan and Buchroithner 2010). Success rate was calculated using 100 training flood locations and prediction of accuracy, using remaining 43 flood locations which were not used during the model building. Class ranges from 'high' to 'very high' susceptibility are considered as potential flood that might occur in future. Success rate and prediction rate are calculated as 0.94 and 0.977, respectively (Table 4). So the prediction accuracy corroborates to 97.7% which validates the FR model used in this flood susceptibility analysis.

Finally, an impact analysis was done to assess the risk factor by probable flood occurrences on the local community and population (Table 5). About 75 villages (20.9%) are situated in the very high flood vulnerable zone and 183 (51.1%) in the high flood vulnerable zone (Fig. 4b). A total of 137292 people are living in those vulnerable zones which require special attention from various levels of governments to take appropriate actions to prevent and mitigate future flood occurrence.

## Conclusion

In the current research, the FR model is used to analyse flood susceptibility zone in the lower part of the Markham river basin (Sub-basin 14). Ten independent conditioning factors, like LULC, elevation, slope, TWI, surface runoff, landform, lithology, distance from the main river, soil texture and soil drainage were derived from the geospatial data sets and used as input into the FR model towards flood prone area mapping. As the result suggests, these ten variables are likely to be major factors to map flood-affected zone. The FR model was used to establish the relationship between past and future flood occurrences. 143 flooded points were generated through "create fishnet" analysis of the PNGRIS national database, field investigation and satellite data before and after flood events. Random sampling technique was used to select 70% input flood data and to calculate frequency ratio for each parameters and their subclasses (Rahmati et al. 2016b). The use of stratified random sampling would be difficult to generate frequency ratio for all classes under each parameter. So random sampling has been mooted by the authors as the prudent choice in this type of studies. FR values for different categories of all conditioning factors were

calculated based on the 100 flood training points (70%). The validation report indicated a higher prediction accuracy of 97.7% which had been enough to validate the FR model that was used for this study. It is obvious that a higher number of input data sets generate higher accuracy. FR model requires a large number of flood points as training like 70% (or 100 points) to generate Frequency Ratio, whereas a less number of input data do not fall under all classes of every parameter. In this point of view, we used 70% flood points for flood map development and 30% (70–30) for the validation process. In case we used 60–40 or 50–50, the results are not same. The accuracy varied 5–10% lower than 70–30 selection method. To validate the superiority, we sought to compare the method used in this study (FR model) with another multi-criteria decision support approach (MCDA) (Samanta et al. 2016a) which was conducted in the same river basin. As per our assessment, FR model produced better results compared to MCDA. This FR model can be used in any other geographical area to develop a flood risk map which can help planners and decision makers to perform proper flood management in future.

**Acknowledgements** The authors are thankful to the PNGUNITECH (Papua New Guinea University of Technology) and to the Department of Surveying and Land Studies for all the facilities made available and availed for the work as a researcher. Satellite digital data available from USGS Global Land Cover Facility and used in this study are also duly acknowledged. The authors gratefully acknowledge the anonymous reviewers for providing their critical comments to improve the quality of this article.

## Compliance with ethical standards

**Conflict of interest** The authors declare that there is no conflict of interest for the publication of this article.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Adiat KAN, Nawawi MNM, Abdullah K (2012) Assessing the accuracy of GIS-based elementary multi criteria decision analysis as a spatial prediction tool—a case of predicting potential zones of sustainable groundwater resources. *J Hydrol* 440–41:75–89
- Aniya M (1985) Landslide-susceptibility mapping in the Amahata river basin, Japan. *Ann Assoc Am Geogr* 75(1):102–114
- Beven and Kirkby (1979) A physically based variable contributing area model of basin hydrology. *Hydrol Sci Bull* 24:43–69
- Billa L, Shattri M, Mahmud AR, Ghazali AH (2006) Comprehensive planning and the role of SDSS in flood disaster management in Malaysia. *Dis Prev Manag* 15:233–240

- Bonham-Carter GF (1994) Geographic information systems for geoscientists: modeling with GIS. In: Bonham-Carter F (ed) Computer methods in the geosciences. Pergamon, Oxford
- Chen YR, Yeh CH, Yu B (2011) Integrated application of the analytic hierarchy process and the geographic information system for flood risk assessment and flood plain management in Taiwan. *Nat Hazards* 59(3):1261–1276
- Chung CF, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. *Nat Hazard* 30:451–472
- Du J, Fang J, Xu W, Shi P (2013) Analysis of dry/wet conditions using the standardized precipitation index and its potential usefulness for drought/flood monitoring in Hunan Province China. *Stoch Env Res Risk Assess* 27(2):377–387
- Haghizadeh A, Siahkamari S, Haghghiabi AH, Rahmati O (2017) Forecasting flood-prone areas using Shannon's entropy model. *J Earth Syst Sci* 126:39. <https://doi.org/10.1007/s12040-017-0819-x>
- Huang X, Tan H, Zhou J, Yang T, Benjamin A, Wen SW, Li S, Liu A, Li X, Fen S, Li X (2008) Flood hazard in Hunan province of China: an economic loss analysis. *Nat Hazards* 47:65–73
- Jebur MN, Mohd Shafri HZ, Pradhan B, Tehrany MS (2013) Per-pixel and object oriented classification methods for mapping urban land cover extraction using SPOT 5 imagery. *Geocarto Int*. <https://doi.org/10.1080/10106049.2013.848944>
- Khan H, Haider S, Saeed K, Ali N (2008) Assessment of potable water quality of Kohat division and its impact on health. *J Chem Soc Pak* 30:246–250
- Kia MB, Pirasteh S, Pradhan B, Rodzi MA, Sulaiman WNA, Moradi A (2012) An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environ Earth Sci* 67:251–264
- Kisi O, Nia AM, Gosheh MG, Tajabadi MRJ, Ahmadi A (2012) Intermittent streamflow forecasting by using several data driven techniques. *Water Resour Manag* 26(2):457–474
- Kugler Z, Groeve TD (2007) The Global Flood Detection system; EUR 23303 EN. Office for Official Publications of the European Communities, Luxembourg, pp 1–45
- Kumar P, Tiwari KN, Pal DK (1991) Establishing SCS runoff curve number from IRS digital database. *J Indian Soc Remote Sens* 19(4):245–251
- Lee MJ, Kang JE, Jeon S (2012) Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. In: Proceedings of Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International. Munich, 895–898
- Li XH, Zhang Q, Shao M, Li YL (2012) A comparison of parameter estimation for distributed hydrological modelling using automatic and manual methods. *Adv Mater Res* 356–360:2372–2375
- Liao X, Carin L (2009) Migratory logistic regression for learning concept drift between two data sets with application to UXO sensing. *IEEE Trans Geosci Remote Sens* 47:1454–1466
- Lohani A, Kumar R, Singh R (2012) Hydrological time series modeling: a comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques. *J Hydrol* 442:23–35
- Manandhar B (2010) Flood Plain Analysis and Risk Assessment of Lothar Khola, MSc Thesis, Tribhuvan University, Phokara, Nepal, pp. 64
- Markantonis V, Meyer V, Lienhoop N (2013) Evaluation of the environmental impacts of extreme floods in the Evros River basin using contingent valuation method. *Nat Hazard* 69:1535–1549
- Merz B, Kreibich H, Schwarze R, Thielen A (2010) Assessment of economic flood damage. *Nat Hazard Earth Syst Sci* 10:1697–1724
- Pal B, Samanta S (2011) Surface runoff estimation and mapping using remote sensing and geographic information system. *Int J Adv Sci Technol* 3(2):106–114
- Pal B, Samanta S, Pal DK (2012) Morphometric and Hydrological analysis and mapping for Watut watershed using Remote Sensing and GIS techniques. *Int J Adv Eng Technol* 2(1):357–368
- Patel DP, Srivastava PK (2013) Flood hazards mitigation analysis using remote sensing and GIS: correspondence with town planning scheme. *Water Resour Manag* 27:2353–2368
- Pourghasemi HR, Pradhan B, Gokceoglu C (2012) Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Nat Hazards* 63:965–996
- Pourghasemi HR, Moradi HR, Aghda SMF, Gokceoglu C, Pradhan B (2014) GIS-based landslide susceptibility mapping with probabilistic likelihood ratio and spatial multi-criteria evaluation models (North of Tehran, Iran). *Arab J Geosci* 7(5):1857–1878
- 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, Buchroithner MF (2010) Comparison and validation of landslide susceptibility maps using an artificial neural network model for three test areas in Malaysia. *Environ Eng Geosci* 16:107–126
- 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
- Qin et al (2011) An approach to computing topographic wetness index based on maximum downslope gradient. *Precision Agric* 12:32–43
- Rahmati O, Zeinivand H, Besharat M (2016a) Flood hazard zoning in Yasooj region, Iran, using GIS and multi-criteria decision analysis. *Geomat Nat Hazard Risk*. <https://doi.org/10.1080/19475705.2015.1045043>
- Rahmati O, Pourghasemi HR, Zeinivand H (2016b) Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto International*. <https://doi.org/10.1080/10106049.2015.1041559>
- Rao KV, Bhattacharya AK, Mishra K (1996) Runoff estimation by curve number method- case studies. *J Soil Water Conserv* 40:1–7
- Regmi NR, Giardino JR, Vitek JD (2010) Modeling susceptibility to landslides using the weight of evidence approach: Western Colorado, USA. *Geomorphology* 115:172–187
- Saha AK, Gupta RP, Sarkar I, Arora KM, Csaplovics E (2005) An approach for GIS-based statistical landslide susceptibility zonation with a case study in the Himalayas. *Landslides* 2(1):61–69
- Samanta S, Pal DK, Lohar D, Pal B (2011) Preparation of digital data sets on land use/land cover, soil and digital elevation model for temperature modelling using Remote Sensing and GIS techniques. *Indian J Sci Technol* 4(6):636–642
- Samanta S, Pal DK, Lohar D, Pal B (2012) Interpolation of climate variables and temperature modeling. *Theor Appl Climatol* 107(1):35–45. <https://doi.org/10.1007/s00704-011-0455-3>
- Samanta S, Koloa C, Pal DK, Palsamanta B (2016a) Flood risk analysis in lower part of Markham River based on multi-criteria decision approach (MCDA). *Hydrology* 3(3):29. <https://doi.org/10.3390/hydrology3030029>
- Samanta S, Koloa C, Pal DK, Palsamanta B (2016b) Estimation of potential soil erosion rate using RUSLE and E30 model. *J Model Earth Syst Environ* 2(3):1–11. <https://doi.org/10.1007/s40808-016-0206-7>
- SCS (1972) Soil Conservation Department, Handbook of Hydrology. Ministry of Agriculture, New Delhi
- Sezer EA, Pradhan B, Gokceoglu C (2011) Manifestation of an adaptive neuro-fuzzy model on landslide susceptibility mapping: Klang Valley Malaysia. *Expert Syst Appl* 38(7):8208–8219

- Solin L (2012) Spatial variability in the flood vulnerability of urban areas in the headwater basins of Slovakia. *Flood Risk Manag* 5:303–320
- Talei A, Chua LHC, Quek C (2010) A novel application of a neuro-fuzzy computational technique in event-based rainfall–runoff modeling. *Expert Syst Appl* 37(12):7456–7468
- Tehrany MS, Pradhan B, Jebur MN (2014a) Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J Hydrol* 512:332–343
- Tehrany MS, Lee MJ, Pradhan B, Jebur MN, Lee S (2014b) Flood susceptibility mapping using integrated bivariate and multivariate statistical models. *Environ. Earth Sci* 72:4001–4015
- Tehrany MS, Pradhan B, Jebur MN (2015) Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stoch Environ Res Risk Assess* 29:1149–1165. <https://doi.org/10.1007/s00477-015-1021-9>
- Tien Bui D, Pradhan B, Lofman O, Revhaug I, Dick OB (2012) Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. *Comput Geosci* 45:199–211
- Tilley JR, Kell RAM, Haro A (2006) Protecting the Markham Bridge, Morobe Province, PNG. In: Proceedings of the 9th International River Symposium, Brisbane, Australia, 4–7 September
- Tiwari MK, Chatterjee C (2010) Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs). *J Hydrol* 382(1):20–33
- Wanders N, Karssen D, de Roo A, de Jong SM, Bierkens MFP (2014) The suitability of remotely sensed soil moisture for improving operational flood forecasting. *Hydrol Earth Syst Sci* 18:2343–2357
- Wang HB, Wu SR, Shi JS, Li B (2013) Qualitative hazard and risk assessment of landslides: a practical framework for a case study in China. *Nat Hazard* 69(3):1281–1294
- WHO (2003) World Health Organization. Disaster Data-key Trends and Statistics in World Disasters Report; WHO: Geneva, Switzerland. [http://www.ifrc.org/PageFiles/89755/2003/43800-WDR2003\\_En.pdf](http://www.ifrc.org/PageFiles/89755/2003/43800-WDR2003_En.pdf) (Accessed 5 April 2017)
- Yalcin A (2008) GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. *CATENA* 72:1–12
- Youssef AM, Pradhan B, Hassan AM (2011) Flash flood risk estimation along the St.Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. *Environ Earth Sci* 62:611–623
- Zou Q, Zhou J, Zhou C, Song L, Guo J (2013) Comprehensive flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy AHP. *Stoch Env Res Risk Assess* 27(2):525–546

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.