Ensemble of data-driven EBF model with knowledge based AHP model for slope failure assessment in GIS using cluster pattern inventory

Biswajeet PRADHAN*, Omar F. ALTHUWAYNEE,
University Putra Malaysia (UPM)
biswajeet24@gmail.com

Commission No.(6811)

Presentation framework

- Introduction
- Motivation
- Study area
- Objectives
- Methodology
- Results and discussion
- Validation
- Conclusion

Introduction

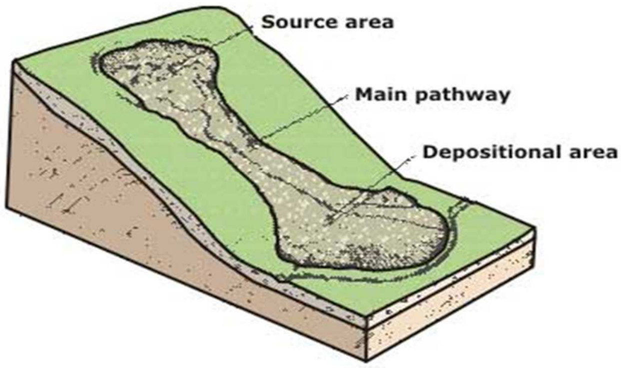
Landslide: Is a downward or outward movement of debris (e.g. soil, rock or vegetation), under the influence of gravity.

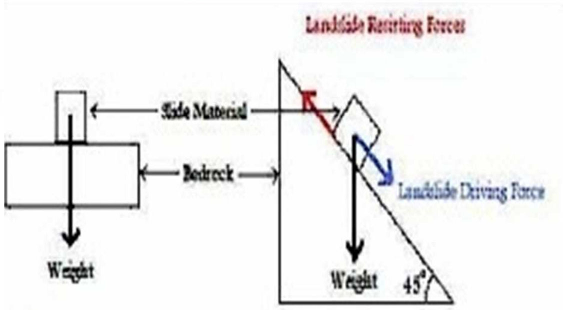
Resisting forces can be significantly reduced in case of rain or earthquake vibrations.

$$F = \frac{\text{Resisting Force(R)}}{\text{Driving Force(D)}}$$

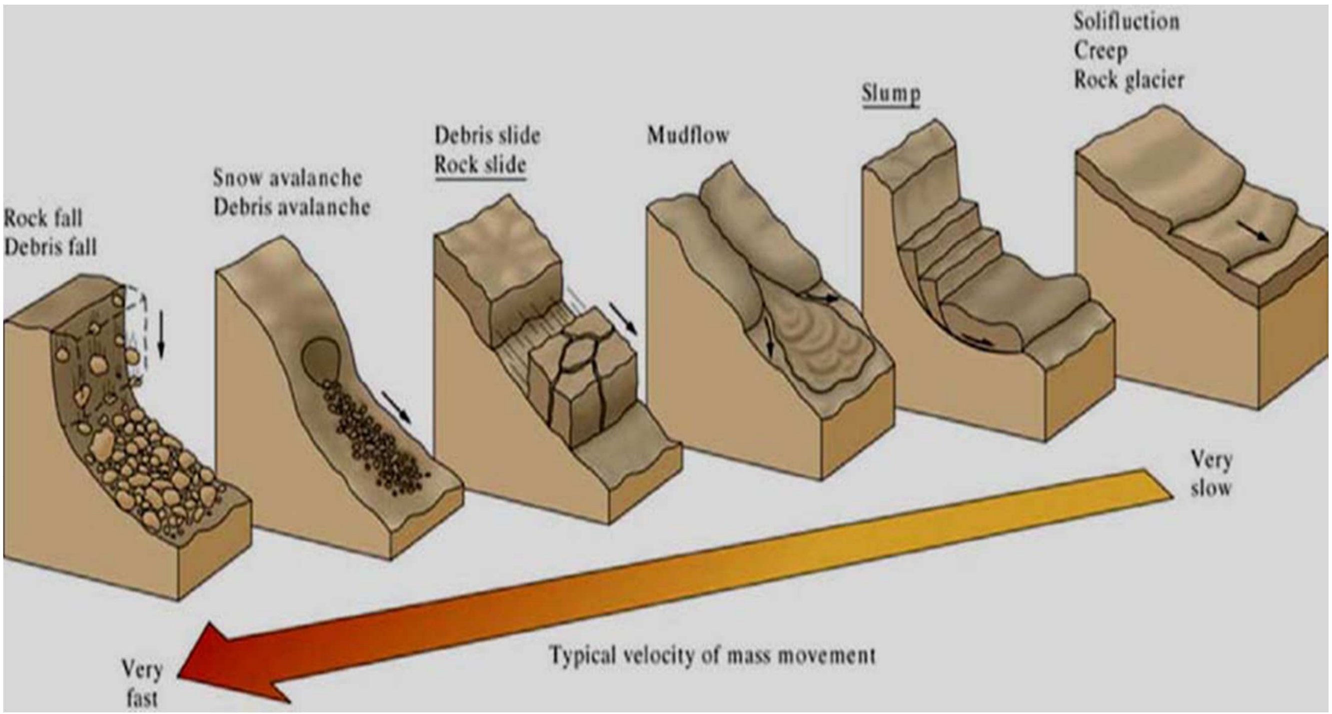
When, $F < 1$ = landslide occur

F: Safety factor





Introduction



Main Factors that cause landslides

1. Slope: The steeper the slope, the larger the threat.
2. Precipitation: Soil is typically more mobile when it is wet.
3. Vegetation: increase stability, reduce water content and control the sediment from eroding down the hill.
4. Soil: most mobile sediments like clay, silt, and mud.
5. Others, elevation, distance from faults and roads.

Motivation

- Enormous property damage, direct and indirect loss of lives (highly urbanized and remote regions) and cost (infrastructure and utilities).
- Retreat the country growth trend.

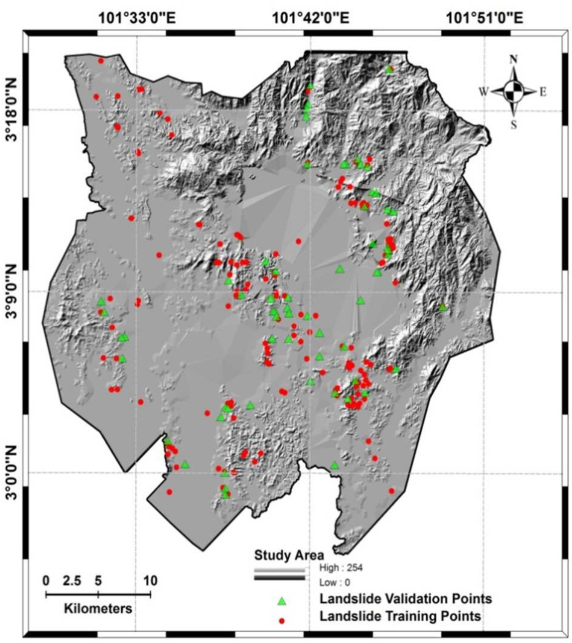
UTM UNIVERSITI TEKNOLOGI MALAYSIA JUPEM MaCGDI
XXV International Federation of Surveyors Congress, Kuala Lumpur, Malaysia, 16 – 2
June 2014
UNIVERSITI TEKNOLOGI MARA UPM UNIVERSITI PUTRA MALAYSIA USM UNIVERSITI SAINS MALAYSIA

Statistics

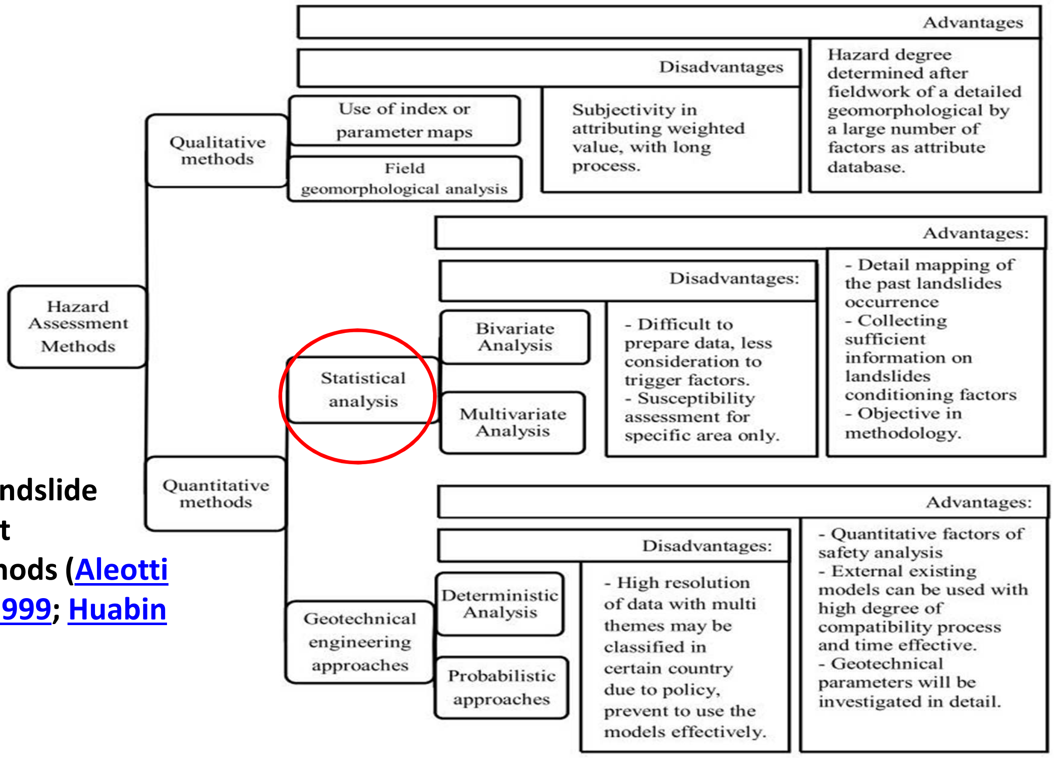
	Type	Date	Location	Notes
1	Road collapsed	29 Nov. 2012	Jalan sungai lalang in kajang	Five people escaped with minor injuries, a car, a van and a motorcycle landed in the ravine
2	Collapse of a concrete embankment	28 Dec. 2012	Bukit setiawangsa, KL.	Residents of 46 houses being evacuated
3	Road collapse	19 Feb. 2013	Ara damansara, Petaling Jaya .	18 families evacuated
4	Soil erosion	27 Mar. 2013	Beringin puchong	A half meter from an apartment
5	Soil erosion	7 May 2013	Bukit gasing, KL.	Nine cars buried , jalan ampang near the scene have been closed to traffic.
6	Soil erosion	9 May 2013	Near Amadesa condominum, KL.	Interrupted the traffic flow

Study area

- Landslide Inventory: (1980-2010) shallow landslides
- Precipitation: Highest amount during Monsoons i.e. 150 to 240 (mm/month)
- Land cover: settlement, peat swamp forest, and abandoned mining, grassland and few shrub areas.
- Temperature: (29 to 32° C).



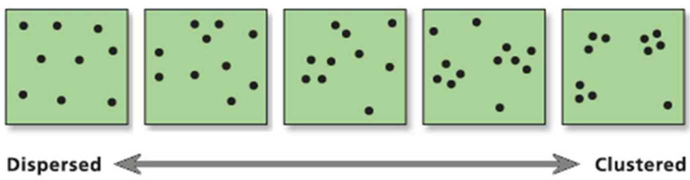
Objectives



Comparisons of landslide hazard assessment classification methods (Aleotti and Chowdhury, 1999; Huabin et al., 2005).

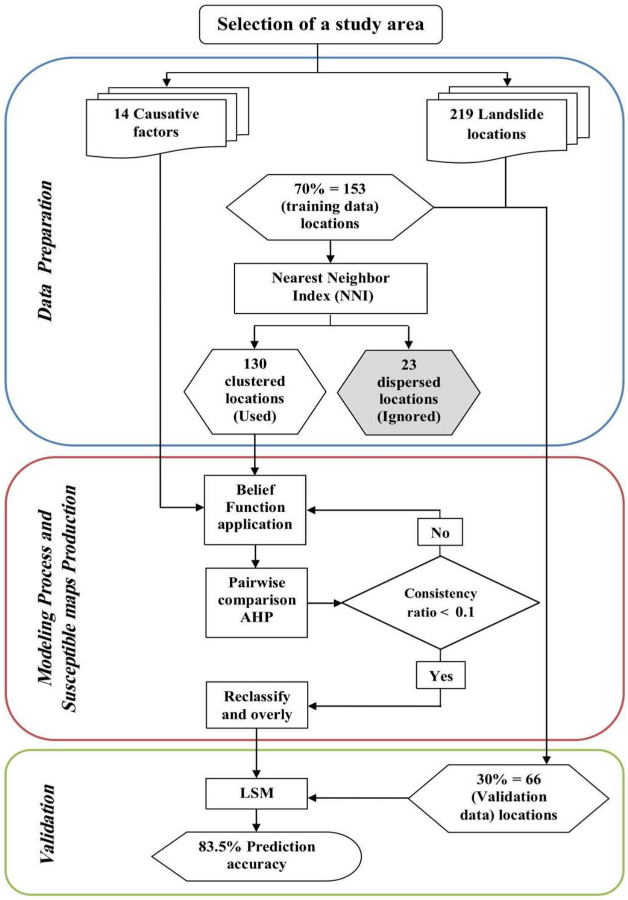
Objectives

1. To test the spatial nature pattern of landslide inventory statistically, i.e. to determine whether it rejects the independency of spatial pattern or not (i.e. random or cluster distribution).



2. To reduce the subjectivity of the experts opinions in AHP model, through developing ensemble quantitative model.

Methodology



Spatial pattern analysis: Nearest Neighbour Index (NNI)

- A 2nd order (local test), mostly describes the overall neighborhood or sub-region patterns (Clark and Evans, 1954).

1. Nearest neighbor distance, which measures the distance from a specific landslide location to all other locations, then, register only the shortest (Eq. 1).

$$\text{Nearest Neighbor Distance} = d(NN) = \sum_{i=1}^{N}\left[\frac{\text{Min}(d_{ij})}{N}\right] \qquad (1)$$

where:

Min (dij): distance between each point and its nearest neighbor (m).
N: number of points.

2. Mean random distance, which measures the expected nearest neighbor distances (i.e. If the spatially random distributed points) (Eq. 2).

$$\text{Mean Random Distance} = d(ran) = 0.5\ \text{SQRT}\left[\frac{A}{N}\right] \qquad (2)$$

where:

A: area of study (m^2)

Spatial pattern analysis: Nearest Neighbour Index (NNI)

3. If the result of NNI (Eq. 3) is <1, it confirm cluster distribution patterns

$$\text{Nearest Neightbor Index} = \text{NNI} = \frac{\text{d(NN)}}{\text{d(ran)}} \quad (3)$$

4. Z-test used to check if the result of Eq. 1 is significantly different from the result of Eq. 2.
   - Negative result of Z-test confirms the cluster nature, and vice versa.

$$Z = \frac{\text{d(NN)} - \text{d(ran)}}{\text{SE}_{\text{d(ran)}}} \quad (4)$$

5. The standard error of the mean random distance is calculated using Eq. 5

$$\text{SE}_{\text{d(ran)}} \approx \text{SQRT}\left[\frac{(4-\pi)A}{4\pi N^2}\right] \approx \frac{0.26136}{\text{SQRT}[N^2/A]} \quad (5)$$

Evidential belief function EBF

The Dempster–Shafer theory of evidence Shafer (1976), considered as a spatial integration model with mathematical representation, mainly used in mineral potential mapping (Carranza, 2009).

Bivariate statically method, with Four output maps:

- Degrees of belief: showed the susceptible areas, Degrees of disbelief: showed the non-susceptible areas, Degrees of uncertainty: showed where the evidences are insufficient to provide the proofs for landslide information, or guide for further field assessment, Degrees of plausibility: represented all the integrated maps evidence except the disbelief map. Generally it shows where spatial evidences are sufficient. Or evidences are inefficient to prove where the landslide triggered factor will effect.

Evidential belief function EBF

mass function. $M: 2^{\Theta} = \{\emptyset, Tp, Tp^{-}, \Theta\}$ $\Theta = \{Tp, Tp^{-}\}$

where: Tp = class pixels effected by landslide
Tp⁻= class pixels not effected landslide

$$\lambda(Tp)Eij = [N(L \cap Eij)/N(L)] \; / \; [N(Eij) - N(L \cap Eij))/(N(A) - N(L)] = N/D$$

$$Bel = \lambda(Tp)Eij \; / \; \sum \lambda(Tp)Eij$$

where
- *N (L∩Eij)*: number of landslide pixels in domain
- *N(L)*: total number of landslide, or $\sum N(L \cap E_{ij})$
- *N(Eij)*: number of pixel in domain
- *N(A)*: total number of pixels in domain, or $\sum N(E_{ij})$
- *N*: proration of landslide occur
- *D*: proportion of non-landslide area

$$\lambda(Tp^{-})Eij = [(N(L) - N(L \cap Eij))/N(L)] \; / \; [(N(A) - N(L) - N(Eij) + N(L \cap Eij))/(N(A) - N(L))] = K/H$$

$$Dis = \lambda(Tp^{-})Eij \; / \; \sum \lambda(Tp^{-})Eij$$

$$Pls = 1 - Dis$$

$$Unc(Ignorance\ or\ doubt) = Pls - Bel$$

$$Unc = 1 - Dis - Bel$$

$$Bel + Unc + Dis = 1$$

where:
K: proportion of landslides that do not occur.
H: proportion of non-landslide areas in other attributes outside class

Weighting of causative factors by AHP integration

The quantified conditioning factors of belief (Bel), acts as the input data for pair-wise analysis instead of classic common 9-point pair-wise rating scale:

1. Predictor rating (PR); Degree (importance).

$$PR = (SA_{max} - SA_{min})/(SA_{max} - SA_{min})_{min}$$

where: SA: Index of spatial association (Bel)

2. Converting the fractional predictor into integer weight.

3. Using consistency ratio (CR<= 0.1): Decision evaluation..

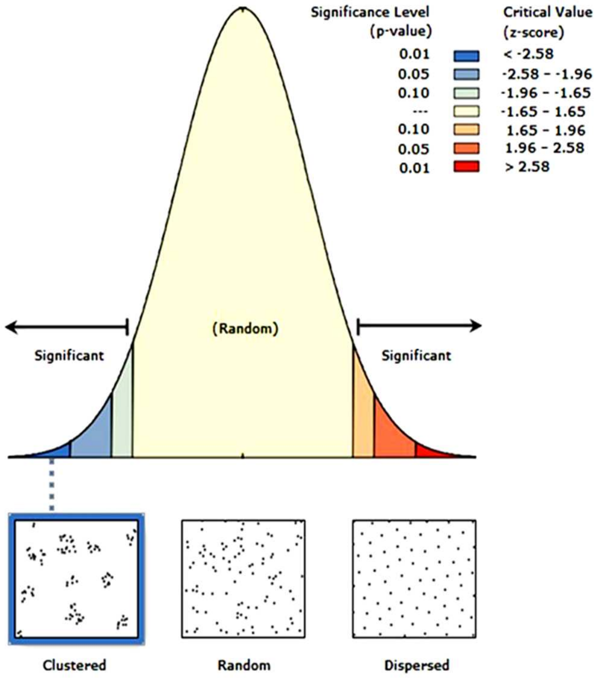
$$CR = CI/RI$$

Where: RI: Average of the resultant consistency index, depends on the order of the given matrix
CI: Consistency index

Results and discussion
NN index results

Expected Mean Distance (1457 m) represents the limit distance which separates between the non-random and random distribution in the current study area.

Observed Mean Distance:	781.195358
Expected Mean Distance:	1457.634980
Nearest Neighbor Ratio:	0.535933
z-score:	-11.017207
p-value:	0.000000



NNI test showed a ratio of 0.53 <1

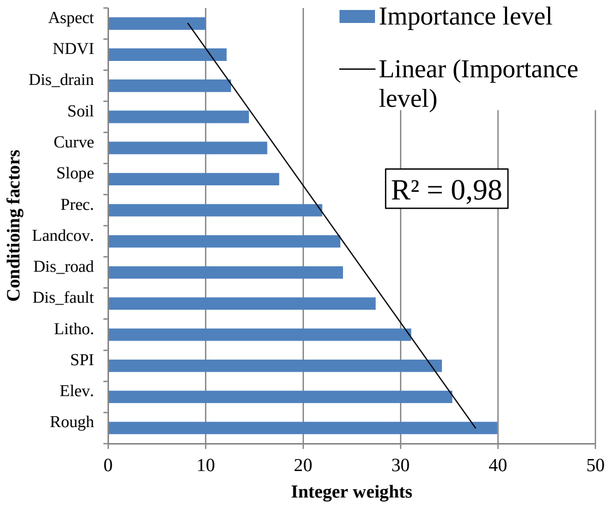
Pairwise comparison results

Table estimated eigenvectors of the pair-wise rating matrix and weights of predictors.

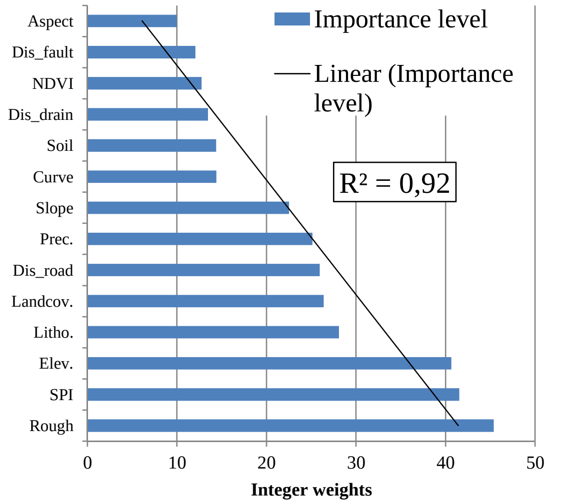
Predictor	Slope	Aspect	Curve	Rough	Elev.	NDVI	SPI	Dis_road	Dis_drain	Litho.	Soil	Landcov.	Prec.	Dis_fault	∑sum	Fractional weight	Integer weight
Slope	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.76	0.055	**18**
Aspect	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.44	0.031	10
Curve	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.71	0.051	16
Rough	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	1.74	0.124	**40**
Elev.	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	1.54	0.110	**35**
NDVI	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.53	0.038	12
SPI	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	1.49	0.107	**34**
Dis_road	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	1.05	0.075	**24**
Dis_drain	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.55	0.039	13
Litho.	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	1.36	0.097	**31**
Soil	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.63	0.045	14
Landcov.	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	1.04	0.074	**24**
Prec.	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.96	0.068	**22**
Dis_fault	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	1.20	0.085	**27**
∑sum	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	14.00	1.00	320.94

Results & Discussion

Conditioning factors of Ensemble modelling using cluster pattern locations



Conditioning factors of Ensemble modelling using random pattern locations

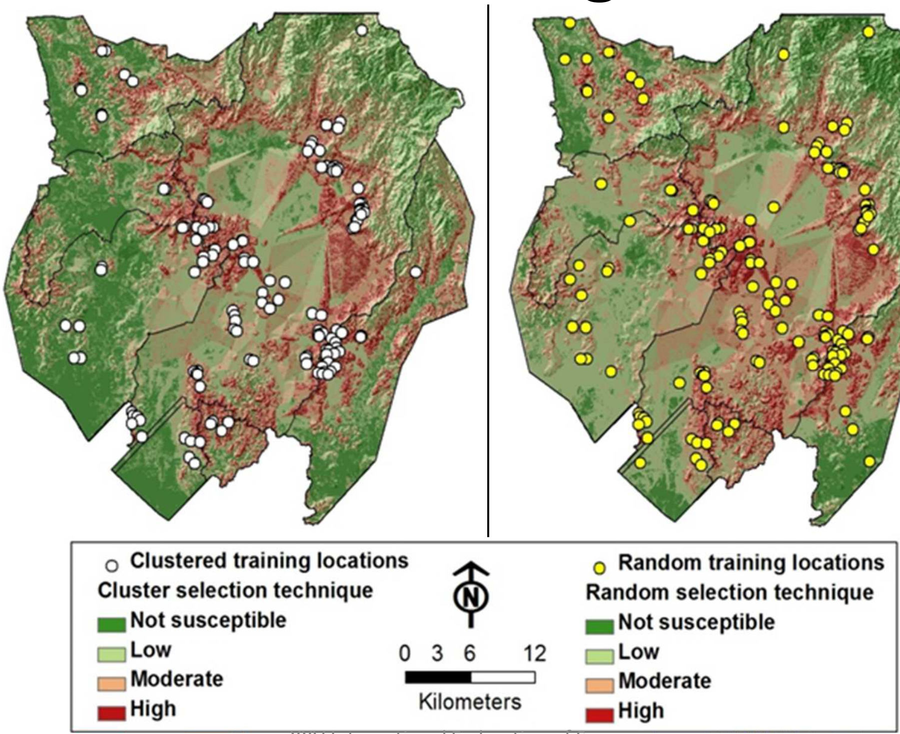


The distance from faults, has a direct relationship with cluster data, as the majority of landslide events accumulate near the fracture faults.

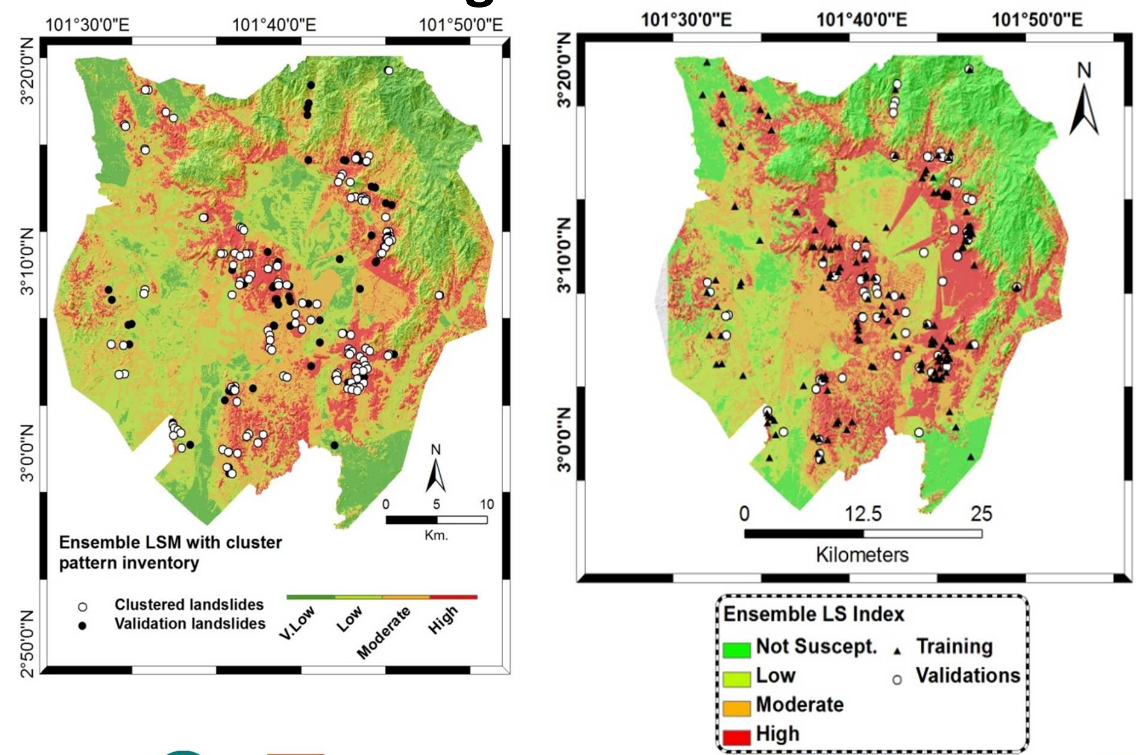
Results & Discussion

1. A total of 219 landslide locations were randomly divided and into 30% (66) validation data, 70% (153) training data.
2. Training data (153) points were tested by NN index.
3. LSM1 using EBF with random pattern locations .
4. LSM2 using EBF with cluster pattern locations.
5. LSM3 using ensemble EBF in pair wise comparison with random pattern locations.
6. LSM4 using ensemble EBF in pair wise comparison with cluster pattern locations.
7. All LSMs results compared, then validated with unused landslide location.

Landslide susceptibility results using EBF



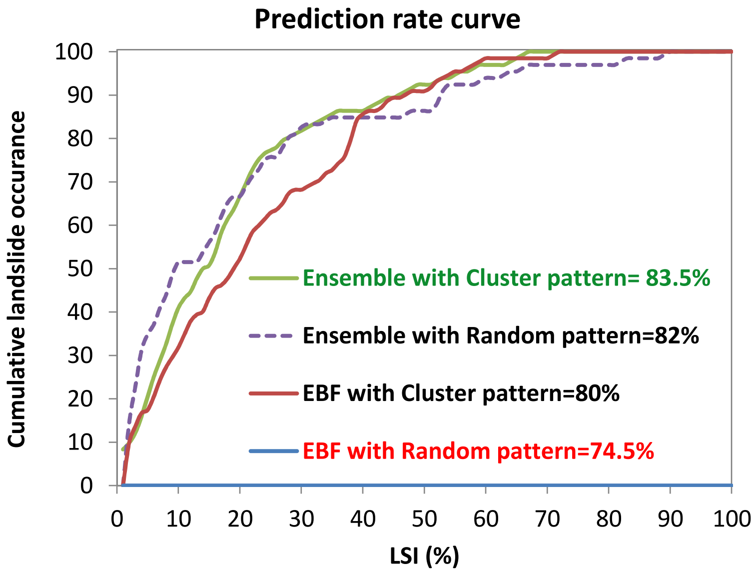
Landslide susceptibility results using Ensemble method



XXV International Federation of Surveyors Congress, Kuala Lumpur, Malaysia, 16 – 21 June 2014

Validation

- Area under prediction curve, plotted with unknown spatial pattern data of 66 landslide locations.



Conclusion

1. Importance of utilizing the computation power of GIS in natural hazards.
2. A 2nd order statistical test of nearest neighbor index was applied to determine whether landslides pattern rejects the independency of spatial pattern or not.
3. Some drawbacks of the AHP and EBF model when applied individually.
4. Landslide inventory shows 88% of events has cluster pattern rather than random pattern of other 12% locations.
5. spatial association between the bivariate EBF and the pair-wise comparison of AHP showed higher prediction accuracy than individual method and in case of cluster pattern than random one.
6. The ensemble optimized the input layers, which can be served as major research advancement in data scarce environments.

Acknowledgment

- UPM-RUGS project grant, vote number: 9344100.
- Financial assistance from Malaysian Land Surveyors Board (LJT).

Publications

Contents lists available at SciVerse ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Application of an evidential belief function model in landslide susceptibility mapping

Omar F. Althuwaynee [a], Biswajeet Pradhan [a,*], Saro Lee [b]

[a] Faculty of Engineering, Spatial and Numerical Modelling Laboratory, Dept. of Civil Engineering, University Putra Malaysia, Serdang, Selangor Darul Ehsan 43400, Malaysia
[b] Korea Institute of Geoscience and Mineral Resources (KIGAM), 92, Gwahang-no, Yuseong-Gu, Daejeon 305-350, South Korea

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena

A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping

Omar F. Althuwaynee [a], Biswajeet Pradhan [a,*], Hyuck-Jin Park [b,*], Jung Hyun Lee [b]

[a] Department of Civil Engineering, Faculty of Engineering, University Putra Malaysia, 43400 Serdang, Selangor Darul Ehsan, Malaysia
[b] Department of Geoinformation Engineering, Sejong University, 98 Gunja-dong, Gwangjin-gu, Seoul 143747, South Korea

Publications

Original Paper

Landslides
DOI 10.1007/s10346-014-0466-0
Received: 13 October 2013
Accepted: 5 January 2014

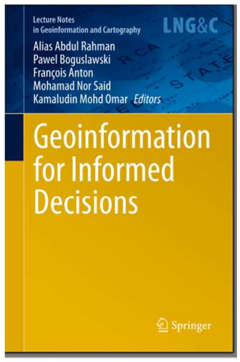
Omar F. Althuwaynee · Biswajeet Pradhan · Hyuck-Jin Park · Jung Hyun Lee

A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping

---

An Alternative Technique for Landslide Inventory Modeling Based on Spatial Pattern Characterization

Omar F. Althuwaynee and Biswajeet Pradhan



Abstract The present study analyses the spatial patterns of historical/present

Thank you for your attention!

University Putra Malaysia
43400 UPM, Serdang
Selangor Darul Ehsan, Malaysia.

Email:
biswajeet24@gmail.com

XXV International Federation of Surveyors Congress, Kuala Lumpur, Malaysia, 16 – 21 June 2014