Determination of Topographic Factors to Initiate Debris Flow Using Statistical Analysis

Chien-Yuan Chen, Ho-Wen Chen, and Zhao-Jun Chen, *Member, IACSIT*

***Abstract*—This study used GIS to determine 19 topographic indexes, four geologic indexes, and two rainfall data derived indexes of the Laonun River Basin, in southern Taiwan. The four topographic factors, including the effective area of basin, elongation ratio of basin, relief energy, and relief volume, were selected among the 19 topographic indexes using the SPSS multi-variable statistical analysis and the principal components analysis. The four topographic factors combine the four geologic factors and two rainfall factors and were selected for estimating debris flow prone creeks. The ten factors were further applied by the Fisher's Discriminant Analysis and Logistic Regression Analysis to evaluate the potentials of debris flow prone creeks in the basin. There were 13 sub-basins initiated debris flows and 41 non-debris flow sub-basins when Typhoon Morakot hit Taiwan in 2009. The validated results show that the correctness of Fisher's model for the samples is 81.48 % and 92.6 % via the Logistic Regression model. Both models showed acceptable accuracy, and the Logistic model had better accuracy herein. The Logistic Regression Analysis was adapted to evaluate the potential of debris flow sub-basins to assist in developing risk management in the basin.**

***Index Terms*—Debris flow, GIS, principal components analysis, linear regression analysis.**

I. Introduction

Typhoon Morakot hit Taiwan in 2009. The typhoon brought over 2,500 mm of torrential rainfall in mountainous southern Taiwan. The rainfall was equivalent to a year of rain in Taiwan and caused serious debris flow hazards in southern Taiwan. Most of the serious debris flow hazards occurred in the Laonun river watershed. An efficient prediction model for debris flow warning is necessary to prevent hazards during climate change-induced torrential rainfall.

The initiation of debris flow is accompanied by abundant losses of slope material in high discharge of steep slope. Topographic indexes are shown to be important in initiating the debris flow. Researchers have proposed numerous represented field conditions of rainfall, geologic, and topographic factors that attributed to the initiation of debris flow [1]–[3].

Investigation of potential debris flow creeks has been studied by numerous researchers that collected related topographic and hydrologic factors to initiate debris flow using statistical analysis. Lee [4] studied rainfall threshold for debris flow warning using non-parametric statistics of the Mann–Whitney–Wilcoxon test (MWW test). He extracted four significant factors to initiate debris flow, including mean streambed slope, effective watershed area, landslide ratio, and lithologic characteristic, for the topographic factors analysis and debris flow potential analysis. Wu [5] studied 34 debris flow potential creeks by Multivariate Statistical Analysis to extract main topographic factors using Principal Component Analysis for watershed area, mean slope, form factor, river density, landslide area, geologic index, effective cumulative rainfall, and effective rainfall intensity. The eight factors were further analyzed by Fisher's Discriminant Analysis for debris flow potential analysis.

Rainfall is the main factor attributed to debris flows in Taiwan. Numerous debris flows are initiated during the peak rainfall intensity of a rainfall event in Taiwan [6], [7]. Separating antecedent rainfall and a rainfall event is required before estimating the critical rainfall factors that initiate debris flow. The relationships between rainfall intensity and duration, intensity and cumulative rainfall, and intensity and antecedent cumulative rainfall are commonly used climate factors to construct a predictive relationship for debris flow warning [6], [8], [9].

II. Study Area and Methodology

A. Study Area

The study area, the Laonun River Basin, is located in Kaoshiun City in southern Taiwan. The Laonun River is a branch of the Gaoping River, the second largest river in Taiwan. Its length is 130 km with a watershed area of 1,370 km2 and average gradient of 1/43. Upstream from the watershed is primarily forest with the South Cross-Island Highway cross the valley. Large areas in the middle and downstream areas have been developed for agricultural activities and for hot spring sightseeing. The greater the development of the areas the more vulnerable they are to natural disasters. Its location is shown in Fig. 1.

The Laonun River Basin is located in the south-west area of the Central Mountain Range. Its topography is higher on the north side and lower on the south side, steeper on the east side and flatter on the west side. The main topographic features are piedmont, meander stream channel, valley, terrace, and alluvial fan. Fig. 2 shows sub-basin division and stream network in the basin.

Manuscript received April 30, 2014; revised June 26, 2014. This work was supported in part by the National Science Council in Taiwan under Grant NSC 99-2625-M-415-003-MY3.

Chien-Yuan Chen and Zhao-Jun Chen are with the Civil and Water Resources Engineering Department, National Chiayi University, Chiayi City 60004, Taiwan R.O.C. (e-mail: chienyuc@mail.ncyu.edu.tw, tainan946012@yahoo.com.tw).

Ho-Wen Chen is with the Department of Environmental Science and Engineering, Tunghai University, Taichung City 40704, Taiwan R.O.C. (e-mail: hwchen@thu.edu.tw).

DOI: 10.7763/IJMLC.2014.V4.471

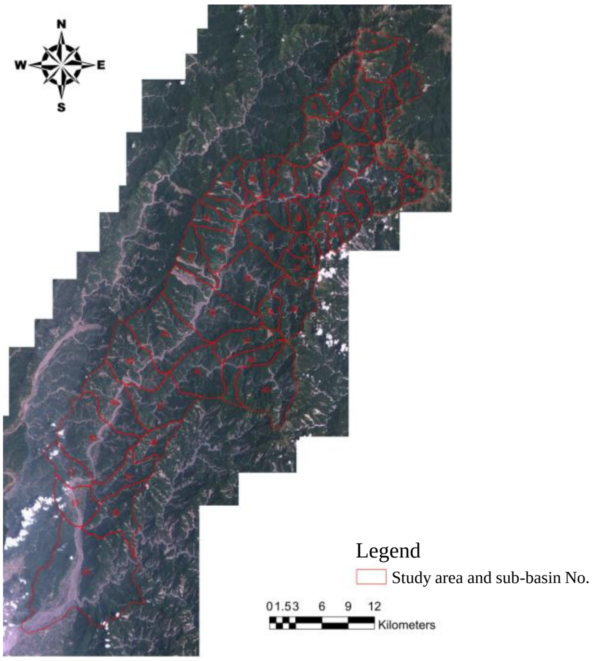

Fig. 1. Study area of the Laonun river basin (satellite image in 2009).

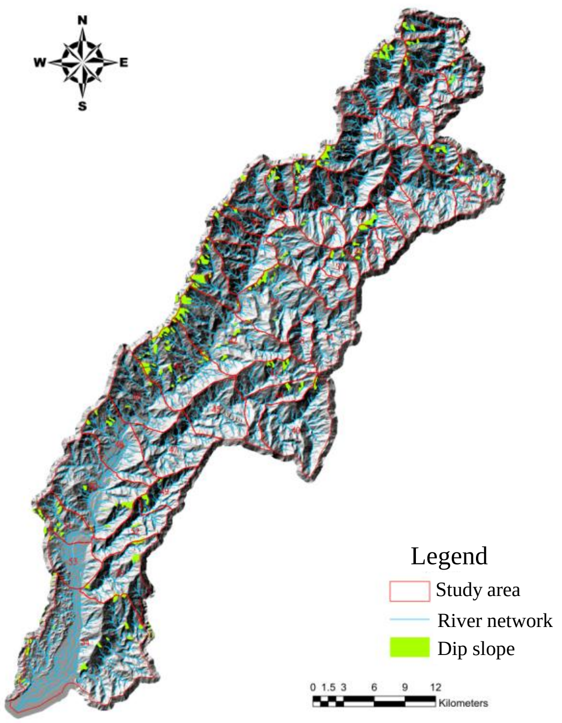

Fig. 2. Basin division, stream network, and dip slope in the study area.

B. Methodology

The study basin is divided into 54 sub-basins to analyze factors that initiate debris flow. Factors that initiate debris flow include climatic, topographic, and geologic conditions. The climatic factors are sourced from rainfall data and topographic factors are calculated from the digital terrain model (DTM) in 40 × 40 m. The regional characteristics that initiate debris flow are analyzed using statistical analysis to extract factors that trigger debris flows and their potential analysis for the studied Laonun River Basin. The purpose of this study is to establish a precise model to locate potential debris flow creeks for hazard prevention and mitigation. The analysis methodology can be divided into four parts:

1) Digital database collection: Digital database include satellite images, DTM, rainfall records, and geology maps are collected for the analysis.
2) Debris flow initiation factors database setting: The study using GIS spatial analysis to estimate geologic and topographic factors to initiate debris flow and collection of rainfall records during 2006 to 2009 for climate factors estimation.
3) Extraction initiation factors of debris flow: The study using Principal Component Analysis to check if the topographic factors have significant differences and show independence. A rainfall event is defined for estimating the climate factors; then, extraction of the main affecting factors, including geologic, topographic, and climatic factors, to determine potential of debris flow for further statistical analysis.
4) Test and validation: Statistical Product and Service Solutions package (SPSS) is used for the multi-variable statistical analysis using Fisher's Discriminant Analysis and Logistic Regression Analysis and cross calibration to set a function for debris flow potential assessment.

TABLE I: DEFINITION OF USED TOPOGRAPHIC FACTORS FOR THE ANALYSIS

Factors	Symbol	Unit	Definition
Effective basin area	A_s	km^2	cell area for slope over 15° in the basin
Perimeter of basin	P	km	the length of a map line that encloses the catchment area of a drainage basin
Basin length	L	km	length in a straight line from the mouth of a stream to the farthest point on the drainage divide of its basin
Length of main channel	L_o	km	the length of the longest succession of segments that connect a source to the outlet of the basin
Total stream length	L_T	km	summation of main stream's and branches' length in the basin
Mean basin width	W	km	the ratio between the area of the basin and the length of the main channel
Sinuousity	J	-	the ratio of channel length and down valley path length
Number of streams	N	-	total numbers of main and branch rivers in the basin
Mean basin elevation	H	m	the mean of difference of maximum and minimum elevation in the basin
Relief energy	R_f	m	maximum difference of elevation
Relief ratio	R_d	-	drop in elevation between the river's source and the river's mouth divided by the total length of the river
Mean basin slope	S	%	the mean of the basin slope
Mean river slope	S_t	%	the ratio of drop in elevation of a stream per unit horizontal distance
Form factor	F	-	the ratio between the area of the basin and the square of the length of the main channel
Compactness	C	-	the ratio between the perimeter of the basin and the diameter of the circle having the same area of the basin
Circularity ratio	M	-	the ratio between the area of the basin and the area of the circle having the same perimeter of the basin
Elongation ratio	E	-	the ratio between the diameter of the circle having the same area of the basin and the length of the main channel
Drainage density	D_s	-	the ratio between the total length of the river network and the basin area
Stream frequency	F_s	-	the ratio between the number of the branches and the area of the basin

III. Spatial Database Setting and Statistical Screening Analysis

Topographic factors of every sub-basin are estimated using spatial analysis in GIS. There are 19 topographic factors in the database including: (1) effective basin area, (2) basin perimeter, (3) basin length, (4) length of main stream, (5) total stream length, (6) mean basin width, (7) sinuousity, (8) number of streams, (9) mean basin elevation, (10) relief energy, (11) relief ratio, (12) mean basin slope, (13) mean river slope, (14) form factor, (15) compactness, (16) circularity ratio, (17) elongation ratio, (18) drainage density, (19) stream frequency (see Table I).

The database of geological factors is sourced from the Central Geological Survey (http://www.moeacgs.gov.tw/) and includes landslide and dip slope area in the basin and Formosa-2 satellite images before and after Typhoon Morakot. The digitalized database includes geological indexes (E_d), fault length, landslide ratio, and ratio of dip slope area (see Table II).

Rainfall isohyets distribution during years 2006 to 2009 showed that downstream area had the higher frequency of debris flows. The results show that high cumulative rainfall is one of the attributing factors to initiate debris flow. The effective cumulative rainfall and effective rainfall intensity are chosen as the climate factors for the analysis (see Table II).

TABLE II: Definition of Geological and Climatic Factors Used for the Analysis

Factors	Symbol	unit	Definition
Geologic index	E_d	km^2	$E_d = (\sum A_i \times W)/A$ A_i: area of rock type i (km^2) A: sub-basin area (km^2) W: weight (unit-free)
Fault length ratio	F_L	km	$F_L = L/P$ L: fault length (km) P: premium of basin (km)
Landslide ratio	G	km^2	G = g/A g: landslide area (km^2) A: basin area (km^2)
Dip slope area ratio	D_A	km^2	$D_A = D_a/A$ D_a: dip slope area (km^2) A: basin area (km^2)
effective cumulative rainfall	R_w	mm	$R_e = \alpha_1 d_1 + \alpha_2 d_2 + \ldots + \alpha_{14} d_{14} = \sum_{t=1}^{14} \alpha_t d_t$ $\alpha_t = 0.5^{t/T}$ α_t = attenuation coefficient, d_t (mm) = daily rainfall of t day, T = half-life, used one day
effective rainfall intensity	I_w	mm/hr	$I_w = R_w / T_w$ T_w = effective rainfall duration by average of the three nearest rain gauge stations during the initiation of debris flow

IV. Results and Discussion

We first used the Kolomogorov-Smironv test to check if the topographic factors show normal distribution. Then, we adopted one of the two methods for "Test for Difference between Means" for the different factors. One-Way ANOVA test (analysis of variance, ANOVA) was used if the factors showed normality; and, the non-parametric Mann-Whitney-Wilcoxon test (MWW test) was used if the factors exhibited non-normality and to find significant effects among the different factors.

A. Kolomogorov-Smironv Test (K-S Test)

The study used the Kolomogorov-Smironv test (K-S test) to test if the 19 topographic factors showed normal distribution (normality). The level of significance in the K-S test was set as 0.01; a level of significance greater than 0.01 is accepted and rejected if smaller than 0.01. The test results show that the three factors for mean basin width (W), mean river slope (S), and form factor (F) showed a level of significance smaller than 0.01 (see Table III). The overall factors showed non-normality for the 19 geologic factors. The MWW test is adopted for the test of difference between means.

TABLE III: K-S Test Data Table

Factors	mean	std dev	min	max	Level of significance
A_s (km)	16.4681	12.2543	3.4427	68.0326	0.400
P (km)	18.9306	7.9454	8.4635	52.3784	0.480
L (km)	5.9840	2.2523	2.5063	15.8583	0.217
L_o (km)	6.3666	2.8442	2.0919	17.7624	0.693
L_T (km)	28.9432	29.0145	5.3540	191.2645	0.190
W (km)	2.9013	2.0247	1.1121	12.7360	<u>0.005</u>
J	1.3681	0.3018	0.4430	2.2423	0.235
N	7.8704	3.4644	2.0000	16.0000	0.498
H (m)	1993.509	724.5209	517.3529	3205.040	0.245
R_f (m)	1643.333	396.0703	720.0000	2380.000	0.972
R_d	0.3396	0.1010	0.1507	0.5448	0.968
S (%)	32.0422	5.1678	14.4800	38.9600	0.053
S_t (%)	1.5461	0.0419	1.3424	1.5684	<u>0.000</u>
F	0.5885	0.7366	0.1829	4.4621	<u>0.000</u>
C	0.7702	0.0772	0.5494	0.9012	0.665
M	0.5990	0.1160	0.3019	0.8122	0.883
E	0.3841	0.0558	0.2730	0.5055	0.988
D_s	0.0015	0.0002	0.0010	0.0022	0.505
F_s	5.57E-07	2.61E-07	1.38E-07	1.15E-06	0.799

B. Non-Parametric Mann-Whitney-Wilcoxon Test (MWW Test)

The MWW test assumed samples show non-normality to compare two samples' differences. The process of the test is to separate the 54 sub-basins into two parts of with and without debris flows initiated by historical records. The basins without debris flow creeks is identified (ID) as 「0」 (validation set), and 「1」 for those with debris flow creeks (training set). The significance level is defined as 0.01, smaller than 0.01 is rejected as no significant difference among these factors. Results of the MWW test (Table IV) for the 19 topographic factors show that the following 12 topographic factors showed significant differences (a significance level greater than 0.01) including effective basin area (A_s), length of main stream (L), width of basin (W), sinuousity (J), relief energy (R_f), relief ratio (R_d), form factor (F), compactness (C), circularity ratio (M), elongation ratio (E), drainage density (D), and stream frequency (F_s).

C. Topographic Factors Screening Analysis Using Principal Component Analysis

There are up to 12 sub-basins' topographic factors that showed significant differences following mutual verification by comparing the K-S test (see Table III) and MWW test (see

Table IV), and therefore need to be further reduced. The study used linear combination of statistical model to screen debris flow potential factors by Factor Analysis in multi-variable statistics of Principal Component Analysis (PCA) and fit Correlation Analysis using a correlation coefficient matrix. The purpose of performing PCA was to screen out the factors that show the largest differences or are independent of others. A Pearson Correlation Matrix was used to standardize and transfer the data from different units of factors. The factor of standardized correlation coefficient is between -1 and +1 and shows the lowest correlation with other factors as the correlation coefficient close to 0, and vice versa as the absolute value close to 1. The selection of principal components follows the Kaiser rule [10] for the factor's corresponding initial eigenvalue (λ) with $\lambda \geqq 1$ and the level of significance is defined as the absolute value of eigenvectors over 0.7 [11]. In general, a higher absolute value of eigenvector shows a close correlation between the two factors and high absolute value of eigenvector of factor is chosen by Pearson correlation coefficient matrix. The order of first principal component is stream frequency, effective basin area, circularity ratio, compactness, and relief ratio by the ranking of the absolute value of eigenvectors (Tables V and VI). It is found that the effective basin area has a higher significant (absolute value of eigenvector) than others and is selected as the first principal component.

TABLE IV: MANN-WHITNEY-WILCOXON TEST DATA TABLE

Factors	ID	No.	mean	Sum	U value	Wilcoxon W	Z value	Level of significance
A_s (km)	0	41	24.3659	999	138	999	-2.5999	0.010
	1	13	37.3846	486				
P (km)	0	41	24.2927	996	135	996	-2.6606	0.008
	1	13	37.6154	489				
L (km)	0	41	23.9268	981	120	981	-2.9640	0.004
	1	13	38.7692	504				
L_0 (km)	0	41	25.7805	1057	196	1057	-1.4264	0.154
	1	13	32.9231	428				
L_T (km)	0	41	23.5854	967	106	967	-3.2473	0.002
	1	13	39.8462	518				
W (km)	0	41	24.4146	1001	140	1001	-2.5594	0.011
	1	13	37.2308	484				
J	0	41	27.3659	1122	261	1122	-0.1113	0.912
	1	13	27.9231	363				
N	0	41	24.1829	991.5	130.5	991.5	-2.7652	0.006
	1	13	37.9615	493.5				
H (m)	0	41	32.1220	1317	77	168	-3.8340	0.001
	1	13	12.9231	168				
R_f	0	41	29.0488	1191	203	294	-1.2851	0.199
	1	13	22.6154	294				
R_d	0	41	30.0488	1232	162	253	-2.1143	0.035
	1	13	19.4615	253				
S (%)	0	41	31.2683	1282	112	203	-3.1261	0.002
	1	13	15.6154	203				
S_L (%)	0	41	31.2927	1283	111	202	-3.1461	0.002
	1	13	15.5385	202				
F	0	41	25.5366	1047	186	1047	-1.6287	0.104
	1	13	33.6923	438				
C	0	41	27.5122	1128	266	357	-0.0101	0.992
	1	13	27.4615	357				
M	0	41	27.5122	1128	266	357	-0.0101	0.992
	1	13	27.4615	357				
E	0	41	27.3415	1121	260	1121	-0.1315	0.896
	1	13	28.0000	364				
D_s	0	41	24.5610	1007	146	1007	-2.4380	0.015
	1	13	36.7692	478				
F_s	0	41	30.0488	1232	162	253	-2.1143	0.035
	1	13	19.4615	253				

The second principal components for absolute value of eigenvectors over 0.7 include effective basin area and mean basin width. The two factors have close correlation, considering the fact that the form factor is more commonly used and is higher in significance than the mean basin width. Form factor is selected as the second principal component. The third principal component is elongation ratio and the fourth principal component is Relief energy. In summary, the four factors for effective basin area, form factor, elongation ratio, and relief energy are chosen as the effective principal components for the following analysis.

TABLE V: INITIAL EIGENVALUE AND TOTAL VARIANCES EXPLAINED TABLE

	Initial eigenvalues			Extraction sums of squared loadings		
No.	total	Variance (%)	Cumulative (%)	total	Variance (%)	Cumulative (%)
1	3.825	31.872	31.872	3.825	31.872	31.872
2	2.496	20.797	52.669	2.496	20.797	52.669
3	1.936	16.133	68.802	1.936	16.133	68.802
4	1.353	11.278	80.080	1.353	11.278	80.080
5	0.885	7.378	87.458	0.885	7.378	87.458
6	0.708	5.904	93.362	0.708	5.904	93.362
7	0.348	2.898	96.260	0.348	2.898	96.260
8	0.246	2.052	98.313	0.246	2.052	98.313
9	0.140	1.169	99.481	0.140	1.169	99.481
10	0.051	0.421	99.903	0.051	0.421	99.903
11	0.010	0.082	99.985	0.010	0.082	99.985
12	0.002	0.015	100.000	0.002	0.015	100.000

TABLE VI: COMPONENT MATRIX AND VALUE OF EIGENVECTOR OF FACTORS IN PRINCIPAL COMPONENT ANALYSIS

	Component			
Factor	1	2	3	4
A_S (km^2)	-0.778	0.148	0.517	0.076
L_0 (km)	-0.677	-0.515	0.352	0.100
W (km)	-0.515	0.796	0.231	-0.123
J	-0.289	-0.691	0.349	0.155
R_f (m)	-0.238	0.255	-0.120	0.817
R_d	0.702	0.238	-0.166	0.452
F	-0.190	0.914	-0.141	-0.111
C	0.727	0.116	0.612	-0.041
M	0.735	0.130	0.598	-0.034
E	0.156	0.257	0.751	0.176
D_s	0.152	-0.024	0.031	-0.608
F_s	0.843	-0.205	-0.101	0.099

There were 10 factors used for the following discriminant and regression analyses for debris flow potential analysis and calculation of its correctness. The selected independent variable includes 4 topographic factors (effective basin area, form factor, elongation ratio, and relief energy), 4 geologic factors (geologic index, fault length, landslide ratio, and ratio of dip slope area), and 2 climatic factors (effective cumulative rainfall and effective rainfall intensity).

The database of 10 factors were selected as dependent variables (outcome variable), and were divided into two categories, the first category 「1」 of 13 sub-basins with records of debris flow and the second category 「0」 of 41 sub-basins without debris flow records.

1) Fisher's discriminant analysis

The results of Fisher's Discriminant Analysis are function coefficient by the classification coefficient of variables (Table VII). The discriminant functions for category 「1」 and category 「0」 are listed as follows:

$$Y_1 = (0.549*A_s) + (1.391*F) + (47.249*E) + (-0.002*R_f) + (-0.923*E_d) + (-124.904*F_L) + (2.949*G) + (42.177*D_A) + (0.068*R_W) + (8.414*I_W) - 142.941 \quad (1)$$

$$Y_0 = (0.318*A_s) + (0.276*F) + (63.941*E) + (0.003*R_f) + (-0.627*E_d) + (-143.363*F_L) + (2.452*G) + (55.851*D_A) + (0.061*R_W) + (7.940*I_W) - 134.263 \quad (2)$$

$$Y = Y_1 - Y_0 = 0.231A_s + 1.115F - 16.692E - 0.005R_f - 0.296E_d + 18.459F_L + 0.497G - 13.674D_A + 0.007R_W - 0.474I_W - 8.678 \quad (3)$$

In the equations, Y_0 is the discriminate function for non-debris flow sub-basins and Y_1 is for debris flow sub-basins. Y is the overall discriminate function for debris flow sub-basins if the value of Y is greater than 0, and for non-debris flow sub-basins if the value is smaller than 0.

A training sample model was adopted to reduce bias (over fitting) and error rate in calculation of the correctness in discriminant analysis in SPSS. The holdout samples method randomly separates the samples into two sets for the training set used for finding the discriminate function of equation and the validation set used for correctness calculation of the discriminate function. The two sets are then changed and the process is repeated for cross validation for the correctness of analysis. The correctness is 88.89 % for the training set, 74.07 % for the validation set, and has overall correctness 81.48 % using Fisher's Discriminant Analysis (see Table VIII).

TABLE VII: COEFFICIENTS OF CLASSIFICATION FUNCTIONS BY FISHER'S DISCRIMINANT ANALYSIS

Factor	category 「0」	category 「1」	coefficient of function (=「1」-「0」)
A_S (km^2)	0.318	0.549	0.231
F	0.276	1.391	1.115
E	63.941	47.249	-16.692
R_f (m)	0.003	-0.002	-0.005
E_d	-0.627	-0.923	-0.296
F_L	-143.363	-124.904	18.459
G	2.452	2.949	0.497
D_A	55.851	42.177	-13.674
R_w (mm)	0.061	0.068	0.007
I_w (mm/hr)	7.940	8.414	0.474
constant	-134.263	-142.941	-8.678

TABLE VIII: RESULTS OF DISCRIMINANT FUNCTION ANALYSIS BY FISHER'S DISCRIMINANT ANALYSIS

Number of training set (54 data)			sum
category	0	1	
0	39 (95.1 %)	2 (4.9 %)	41 (100 %)
1	4 (30.8 %)	9 (69.2 %)	13 (100 %)
correctness: [(39+9)/(41+13)] ×100 % = 88.89 %			
Number of validation set (54 data)			sum
category	0	1	
0	32 (78 %)	9 (22 %)	41 (100 %)
1	5 (38.5 %)	8 (61.5 %)	13 (100 %)
correctness: [(32+8)/(41+13)] ×100 % = 74.07 %			
Number of overall set (108 data)			sum
category	0	1	
0	71 (86.6 %)	11 (13.4 %)	82 (100 %)
1	9 (34.6 %)	17 (65.4 %)	26 (100 %)
correctness: [(71+17)/(82+26)] ×100 % = 81.48 %			

2) Logistic regression analysis

The regression function for debris flow potential analysis was further studied by Logistic Regression Analysis using the ten selected factors. The coefficients of function were obtained from logistic coefficient (B) in classification table of Logistic Regression Analysis (see Table IX). The regression equation was further examined by the Chi-square Test that was significantly (p-value) smaller than 0.05 with acceptable goodness of fit in 95 % of confidence interval. The regression function is listed as follows:

$$Y = 0.165A_s + 0.817F - 14.822E - 0.003R_f - 0.473E_d + 14.150F_L + 0.490G - 29.906D_A + 0.004R_W + 0.347I_W - 3.653 \quad (4)$$

Results of the Logistic Regression Analysis show that the correctness is 100 % for non-debris flow sub-basins (category 「0」), 69.2 % for debris flow sub-basins (category「1」), with 92.6 % of overall correctness.

TABLE IX: CLASSIFICATION COEFFICIENT TABLE BY LOGISTIC REGRESSION ANALYSIS

Factor	B	S.E.	Wald	degrees of freedom	p-value	Exp (B)
A_S (km^2)	0.165	0.084	3.827	1	0.050	1.179
F	0.817	0.595	1.881	1	0.170	2.263
E	-14.822	12.049	1.513	1	0.219	0.000
R_f (m)	-0.003	0.002	2.064	1	0.151	0.997
E_d	-0.473	0.574	0.681	1	0.409	0.623
F_L	14.150	20.144	0.493	1	0.482	1397007
G	0.490	0.392	1.565	1	0.211	1.632
D_A	-29.906	22.055	1.839	1	0.175	0.000
R_w (mm)	0.004	0.004	1.238	1	0.266	1.004
I_w (mm/hr)	0.347	0.360	0.931	1	0.335	1.415
constant	-3.653	8.240	0.196	1	0.658	0.026

3) Comparisons of the two analyzed methodologies

The correctness by Fisher's model is 69.23 % for debris flow sub-basins and 95.12 % for non-debris flow sub-basins, with an overall correctness of 81.48 %. The correctness is 69.23 % for debris flow sub-basins and 100 % for non-debris flow sub-basins, with an overall correctness of 92.6 % by the Logistic Regression model. Fisher's model shows the same correctness for Logistic Regression model in category 「1」(debris flow sub-basin); and lower than Logistic Regression model in category「0」(non-debris flow sub-basin). In general, the prediction model by Logistic Regression model has greater correctness than by Fisher's model. The study suggests using the regression function by Logistic Regression model for potential debris flow analysis and is expected to apply to other areas.

V. CONCLUSION

The study used Fisher's Discriminant Analysis and Logistic Regression Analysis for evaluation of debris flow potentials and its correctness by setting up a function of linear equation. The following list summarizes the results:

1) There are 54 sub-basins, including 13 with and 41 without debris flow records, in the study area. Screening analysis for 19 topographic factors were analyzed using the Kolomogorov-Smironv test (K-S test), Mann-Whitney-Wilcoxon test (M-W-W test), Correlation Analysis, and Principal Component Analysis. Four main topographic factors, including effective watershed area, form factor, elongation ratio, and relief energy, were selected for the analysis.
2) Ten factors showed significant effects to the initiation of debris flow: Effective basin area, form factor, elongation ratio, relief energy, geologic index, fault length, landslide ratio, ratio of dip slope area, effective cumulative rainfall, and effective rainfall intensity.
3) The correctness for the debris flow sub-basins was 88.89 % and 74.07 % for the none-debris flow sub-basins. The overall correctness for the two sets was 81.48 % using Fisher's Discriminant Analysis.
4) The correctness for the debris flow sub-basins was 69.23 % and 100.0 % for the none-debris flow sub-basins using Logistic Regression Analysis. The overall correctness for the two sets was 92.6 % using Logistic Regression Analysis.
5) The study suggests using Logistic Regression Analysis for debris flow potential analysis in the study basin. The regression equation was set as: $Y = 0.165A_s + 0.817F - 14.822E - 0.003R_f - 0.473E_d + 14.150F_L + 0.490G - 29.906D_A + 0.004R_w + 0.347I_w - 3.653$.

References

[1] P. S. Lin, J. Y. Lin, J. C. Hung, and M. D. Yang, "Assessing debris-flow hazard in a watershed in Taiwan," *Engineering Geology*, vol. 66, pp. 295-313, 2002.

[2] P. S. Lin, J. Y. Lin, S. Y. Lin, and J. Lai, "Hazard assessment of debris flows by statistical analysis and GIS in central Taiwan," *International Journal of Applied Science and Engineering*, vol. 4, no. 2, pp. 165-187, 2006.

[3] C. Y. Chen and F. C. Yu, "Morphometric analysis of debris flows and their source areas using GIS," *Geomorphology*, vol. 129, pp. 387-397, 2011.

[4] M. H. Lee, "A rainfall-based debris-flow warning analysis and its application," Ph.D. dissertation, Department of Hydraulics & Ocean Engineering, National Cheng Kung University, Tainan, Taiwan R.O.C., 2006.

[5] W. H. Wu, "Debris flow potential analysis and outflow debris masses estimation," M.S. thesis, Department of Construction Engineering, Chaoyang University of Technology, Taichung, Taiwan R.O.C., 2006.

[6] C. Y. Chen, T. C. Chen, F. C. Yu, W. H. Yu, and C. C. Tseng, "Rainfall duration and debris-flow initiated studies for real-time monitoring," *Environmental Geology*, vol. 47, no. 5, pp. 715-724, 2005.

[7] C. C. Huang, "Gaussian-distribution-based hyetographs and their relationships with debris flow initiation," *Journal of Hydrology*, vol. 411, pp. 251-265, 2011.

[8] S. H. Cannon and S. D. Ellen, "Rainfall conditions for abundant debris avalanches, San Francisco Bay region," *California Geology*, vol. 38, no. 12, pp. 267-272, 1985.

[9] G. Crosta and P. Frattini, "Rainfall thresholds for triggering soil slips and debris flow," in *Proc. the 2nd Plinius International Conference on Mediterranean Storms*, Siena, Italy, 2001, pp. 463-488.

[10] H. F. Kaiser, "The application of electronic computers to factor analysis," *Educational and Pshchological Mesurement*, vol. 20, pp. 141-151, 1960.

[11] S. E. Pratsinis, M. D. Zeldin, and E. C. Ellis, "Source resolution of the fine carbonaceous aerosol by principal component-stepwise regression analysis," *Environ. Sci. Technol.*, vol. 22, pp. 212-216, 1988.

Chien-Yuan Chen was born on October 27, 1969, in Changhua City, Taiwan. He got his Ph.D. in civil & environmental engineering, University of Southern California, Los Angeles, USA in 2001. M.Sc. in civil engineering, National Cheng Kung University, Taiwan in 1995.

He was an associate research fellow (2001-2006) at Slopeland Disaster Reduction Division, National Science and Technology Center for Disaster Reduction (NCDR), Sindian District, New Taipei City, Taiwan. He was an assistant professor (2006-2009) and associate professor (2009-2013) at the Department of Civil and Water Resources Engineering, National Chiayi University. He is currently as a professor (since 2013) at the Department of Civil and Water Resources Engineering, National Chiayi University, Chiayi City, Taiwan. His research is focused on the disaster prevention management and system, disaster prevention education, debris flow and landslide hazards prevention and mitigation, geotechincal engineering, GIS application and numerical modeling.

Prof. Chen is a member of Disaster Management Society of Taiwan.

Ho-Wen Chen was born on February 4, 1970, in Changhua City, Taiwan. He got his Ph.D. in National Cheng Kung Univ, Taiwan in 2000. M.Sc. in environmental engineering, National Cheng Kung University, Taiwan, 1995. He was an assistant professor (2001-2006) and associate professor (2006-2010) at Department of Environmental Engineering and Management, Chaoyang University of Technology. He is currently as a professor (since 2010) at the Department of Environmental Science and Engineering, Tunghai University, Taichung City, Taiwan. His research is focused on sustainable system engineering, artificial intelligent and data mining, environmental informatics, water resources management, environmental forensics.

Zhao-Jun Chen was born on February 4, 1963, in Chiayi City, Taiwan. He received his M.Sc. in civil and water resources engineering, National Chiayi University. Taiwan, 2011. He is currently as an engineering in engineering consultant company.