An ensemble approach of feature selection and machine learning models for regional landslide susceptibility mapping in the arid mountainous terrain of southern Peru

Chandan Kumar[1*], Gabriel Walton[1], Paul Santi[1], Carlos Luza[2]

[1]Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO, 80401, USA

[2]Department of Geology, Geophysics and Mines, Universidad Nacional de San Agustín, Arequipa, 04000, Peru.

*Correspondence: ckumar@mtu.edu (Chandan Kumar)

Supplementary materials

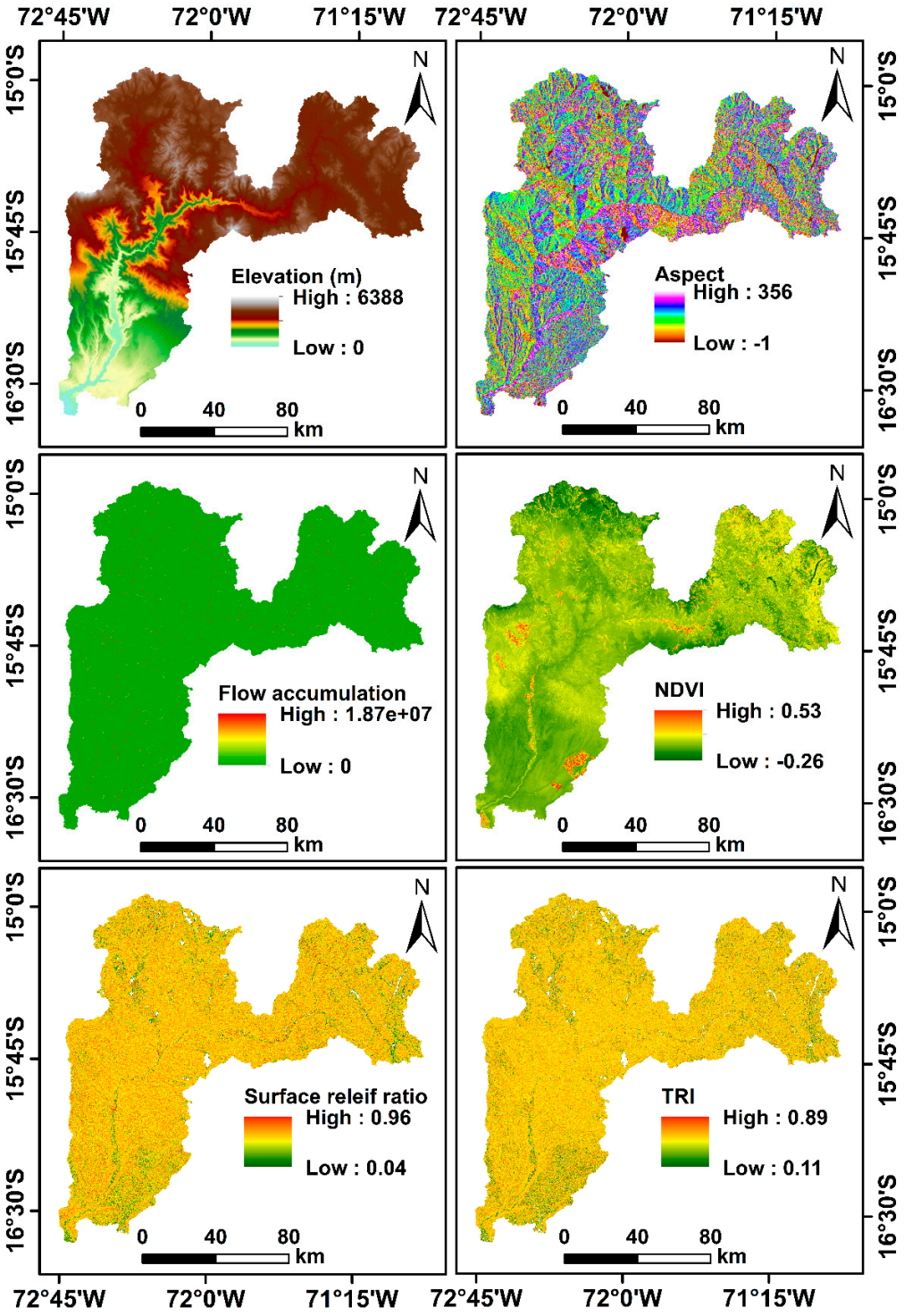


Figure S1. (Continue on next page).

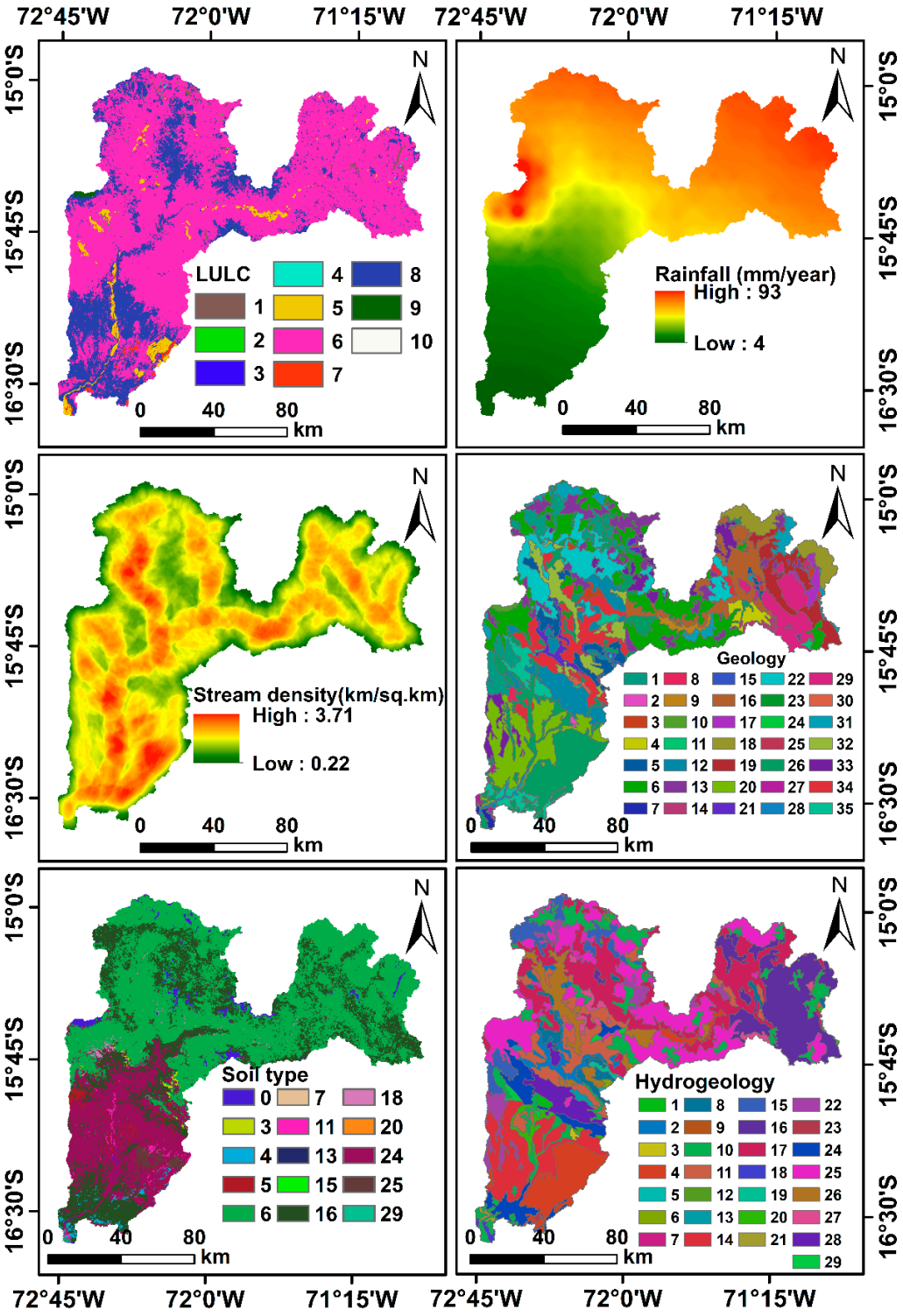


Figure S1. (Continue on next page).

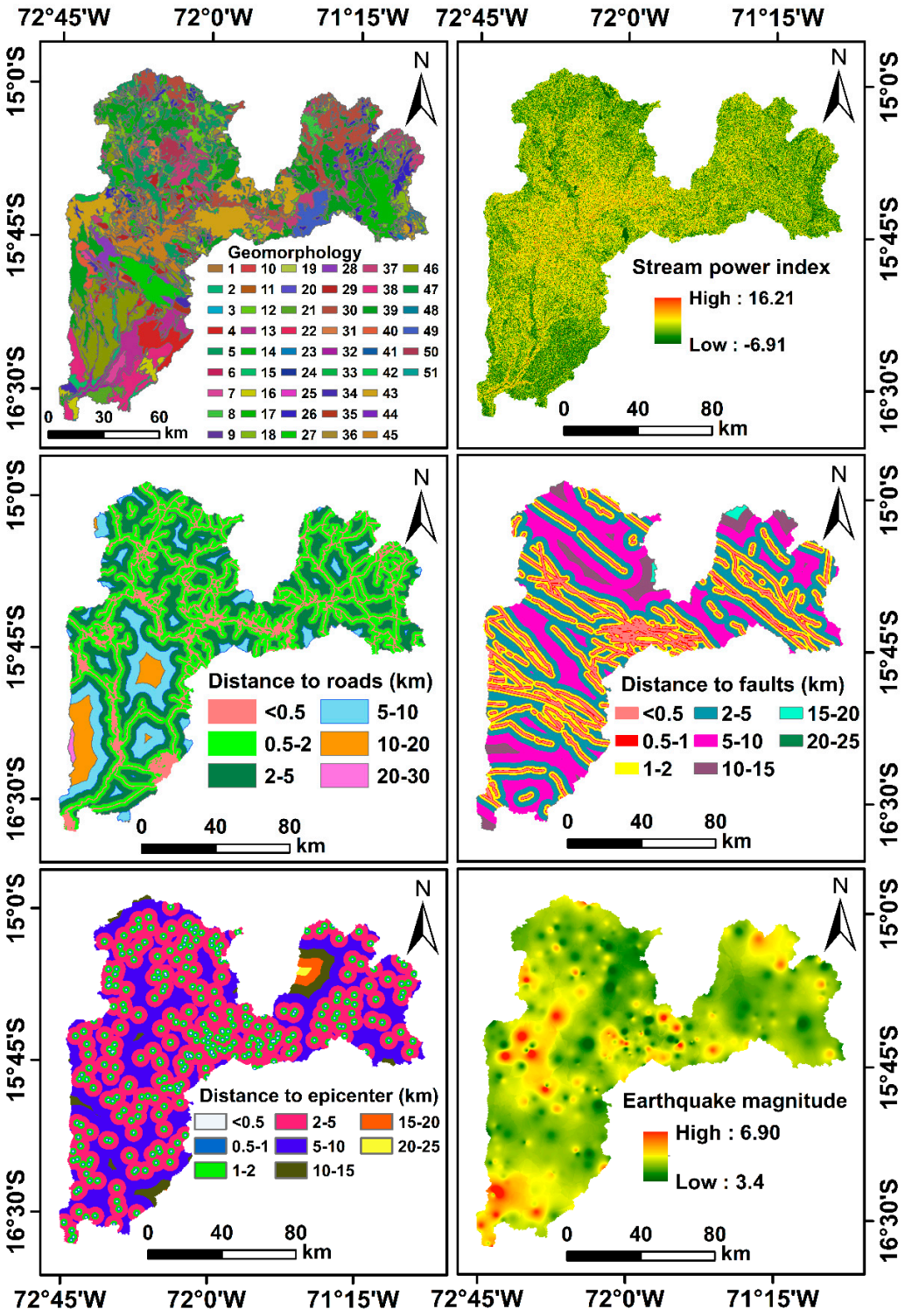


Figure S1. Landslide Influencing Factors (LIFs). The nomenclature of legends is given in Table S1.

Table S1. Frequency Ratio (FR) analysis using historical landslides and Landslide Influencing Factors (LIFs).

LIFs	Class	No. of Landslides	% of Landslides	Class pixels	% of Class pixels	FR
Elevation (m)	1. 0 – 794	100	0.068	480769	0.026	2.677
	2. 794 – 1408	92	0.063	2267987	0.120	0.522
	3. 1408 – 2075	179	0.122	1493053	0.079	1.543
	4. 2075 – 2820	186	0.127	992039	0.053	2.413
	5. 2820 – 3503	133	0.091	1229452	0.065	1.392
	6. 3503 – 4074	138	0.094	1881709	0.100	0.944
	7. 4074 – 4488	259	0.177	3241791	0.172	1.028
	8. 4488 – 4810	150	0.102	4790689	0.254	0.403
	9. 4810 – 6388	227	0.155	2460958	0.131	1.190
Aspect	1. Flat (-1)	0	0.000	398419	0.021	0.000
	2. North (0 – 22.5)	29	0.020	903399	0.048	0.412
	3. Northeast (22.5 – 67.5)	88	0.060	2116190	0.113	0.534
	4. East (67.5 – 112.5)	123	0.084	2298048	0.122	0.687
	5. Southeast (112.5–157.5)	196	0.134	2460268	0.131	1.023
	6. South (157.5 – 202.5)	306	0.209	2489875	0.132	1.578
	7. Southwest (202.5 – 247.5)	340	0.232	2630943	0.140	1.660
	8. West (247.5 – 292.5)	206	0.141	2424781	0.129	1.091
	9. Northwest (292.5 – 337.5)	126	0.086	2178189	0.116	0.743
	10. North (337.5 – 360)	50	0.034	902844	0.048	0.711
Slope (degree)	1. 0 – 5.36	18	0.012	4161103	0.221	0.056
	2. 5.36 – 10.72	47	0.032	4221205	0.224	0.143
	3. 10.72 – 16.38	92	0.063	3278096	0.174	0.361
	4. 16.38 – 22.34	204	0.140	2451521	0.130	1.070
	5. 22.34 – 28.89	328	0.224	1949884	0.104	2.163
	6. 28.89 – 35.74	363	0.248	1495083	0.080	3.123
	7. 35.74 – 44.38	305	0.209	974108	0.052	4.027
	8. 44.38 – 75.95	107	0.073	271956	0.014	5.060
Profile curvature	1. -19.31 – -3.04	11	0.008	64443	0.003	2.196
	2. -3.04 – -1.58	143	0.098	704595	0.037	2.612
	3. -1.58 – -0.93	175	0.120	1761407	0.094	1.278

	4. -0.93 – -0.44	329	0.225	5032020	0.267	0.841
	5. -0.44 – 0.05	356	0.243	7004930	0.372	0.654
	6. 0.05 – 0.54	366	0.250	3814299	0.202	1.235
	7. 0.54 – 1.68	77	0.053	438771	0.023	2.258
	8. 1.68 – 22.18	7	0.005	17982	0.001	5.009
Topographical Position Index (TPI)	1. -69.44 – -3.73	87	0.060	648192	0.034	1.725
	2. -3.73 – -1.42	217	0.148	3241093	0.173	0.860
	3. -1.42 – 0.31	300	0.205	7140337	0.380	0.540
	4. 0.31 – 2.61	445	0.304	6123376	0.326	0.934
	5. 2.61 – 77.56	415	0.284	1635487	0.087	3.261
Topographical Roughness Index (TRI)	1. 0.11 – 0.30	6	0.004	851615	0.046	0.089
	2. 0.30 – 0.42	77	0.053	2812766	0.151	0.347
	3. 0.42 – 0.50	415	0.283	6683222	0.360	0.788
	4. 0.50 – 0.58	652	0.445	6001420	0.323	1.379
	5. 0.58 – 0.89	314	0.214	2229689	0.120	1.787
Topographical Wetness Index (TWI)	1. 2.09 – 5.62	1022	0.699	6625967	0.352	1.984
	2. 5.62 – 7.48	352	0.241	8033621	0.427	0.564
	3. 7.48 – 10.22	79	0.054	2698449	0.144	0.377
	4. 10.22 – 14.03	11	0.008	1110927	0.059	0.127
	5. 14.03 – 27.05	0	0.000	333992	0.018	0.000
Surface Relief Ratio (SRR)	1. 0.04 – 0.28	3	0.002	910731	0.049	0.042
	2. 0.28 – 0.41	56	0.038	3055254	0.164	0.234
	3. 0.41 – 0.50	323	0.222	6343187	0.341	0.650
	4. 0.50 – 0.58	633	0.434	5920525	0.318	1.364
	5. 0.58 – 0.96	449	0.308	2364319	0.127	2.424
Rainfall (mm/day)	1. 3.80 – 11.87	326	0.223	3766110	0.200	1.114
	2. 11.87 – 22.04	224	0.153	2207431	0.117	1.306
	3. 22.04 – 34.32	178	0.122	1772662	0.094	1.292
	4. 34.32 – 43.79	333	0.227	2393100	0.127	1.791
	5. 43.79 – 50.10	199	0.136	2986308	0.159	0.858
	6. 50.10 – 55.71	119	0.081	3517894	0.187	0.435
	7. 55.71 – 64.48	84	0.057	2120775	0.113	0.510
	8. 64. 48 – 93.24	1	0.001	76279	0.004	0.169
Stream density km/km^2	1. 0.22 – 1.26	88	0.060	1839467	0.098	0.616
	2. 1.26 – 1.78	262	0.179	4352069	0.231	0.775
	3. 1.78 – 2.26	331	0.226	5396161	0.286	0.789
	4. 2.26 – 2.76	491	0.335	5090260	0.270	1.241
	5. 2.76 – 3.71	292	0.199	2160527	0.115	1.739

Direct radiation (WH/m^2)	1. 107512 – 415982	53	0.036	128739	0.007	5.297
	2. 415982 – 496453	155	0.106	380292	0.020	5.245
	3. 496453 – 556805	200	0.137	714417	0.038	3.602
	4. 556805 – 607100	223	0.152	1420235	0.075	2.020
	5. 607100 – 657394	230	0.157	2903352	0.154	1.019
	6. 657394 – 717747	229	0.156	1757987	0.093	1.676
	7. 717747 – 778099	186	0.127	2255227	0.120	1.061
	8. 778099 – 828393	129	0.088	3832186	0.203	0.433
	9. 828393 – 962511	59	0.040	5446012	0.289	0.139
Direct Duration Radiation (DDR) (H)	1. 382 – 1128	10	0.007	78889	0.004	1.631
	2. 1128 – 1266	48	0.033	279861	0.015	2.207
	3. 1266 – 1363	117	0.080	620434	0.033	2.427
	4. 1363 – 1444	167	0.114	1041565	0.055	2.063
	5. 1444 – 1518	230	0.157	1560969	0.083	1.896
	6. 1518 – 1593	263	0.180	2258604	0.120	1.498
	7. 1593 – 1668	295	0.202	3306968	0.176	1.148
	8. 1668 – 1742	221	0.151	4673482	0.248	0.608
	9. 1742 – 1846	113	0.077	5017675	0.266	0.290
Normalized Vegetation Index (NDVI)	1. -0.26 – -0.03	114	0.078	1403446	0.074	1.045
	2. -0.03 – 0.03	722	0.493	6900574	0.366	1.346
	3. 0.03 – 0.08	558	0.381	7918335	0.420	0.907
	4. 0.08 – 0.16	67	0.046	2203094	0.117	0.391
	5. 0.16 – 0.53	3	0.002	415110	0.022	0.093
Geology	1. Alpabamba formation	30	0.020	735365	0.039	0.525
	2. Ambo group	6	0.004	16841	0.001	4.584
	3. Andesite, rhyolite, riodacite, dacite	13	0.009	86865	0.005	1.926
	4. Arca formation, Picotani	18	0.012	208337	0.011	1.112
	5. Arcurquina formation	21	0.014	482988	0.026	0.559
	6. Barroso group	191	0.130	2395006	0.127	1.026
	7. Camana formation	17	0.012	56772	0.003	3.853

	8. Castrovirreyna group	3	0.002	36202	0.002	1.066
	9. Colluvial deposit	43	0.029	111783	0.006	4.950
	10. Diorite, granodiorite	4	0.003	41675	0.002	1.235
	11. Fluvioglacial deposits	0	0.000	21782	0.001	0.000
	12. Gabrodiorite, tonalite, granite, diorite, microdiorite	146	0.100	1034885	0.055	1.815
	13. Glacial deposit	16	0.011	1089320	0.058	0.189
	14. Granodiorite, dacite, andesite, diorite, tonalite	3	0.002	70284	0.004	0.549
	15. Huaylillas formation	9	0.006	29808	0.002	3.885
	16. Ichicollo formation	103	0.070	1009950	0.054	1.312
	17. Ipururo formation	0	0.000	157120	0.008	0.000
	18. Marine deposit	26	0.018	566048	0.030	0.591
	19. Maure formation	55	0.038	756985	0.040	0.935
	20. Moquegua formation	161	0.110	1379283	0.073	1.502
	21. Murco formation	8	0.005	244654	0.013	0.421
	22. Orcopampa formation	120	0.082	1389260	0.074	1.111
	23. Palca group	0	0.000	7236	0.000	0.000
	24. Para formation	0	0.000	8570	0.000	0.000
	25. Puno group	153	0.105	2495363	0.132	0.789
	26. Recent alluvial deposit	7	0.005	52862	0.003	1.704
	27. Recent eluvial deposit	13	0.009	782251	0.042	0.214
	28. River deposit	1	0.001	34921	0.002	0.368
	29. Sillapaca group	31	0.021	483047	0.026	0.826
	30. Socosani formation	18	0.012	624865	0.033	0.371
	31. Tacaza group	53	0.036	621034	0.033	1.098
	32. Volcanic Astobamba	130	0.089	1205856	0.064	1.387

	33. Volcanic Senca	65	0.044	601266	0.032	1.391
Geomorphology	1. Alluvial high terrace	33	0.023	62536	0.003	6.790
	2. Alluvial terrace	6	0.004	146542	0.008	0.527
	3. Alluvial plain	0	0.000	2873	0.000	0.000
	4. Aluvio-torrential slope or foothills	5	0.003	588991	0.031	0.109
	5. Basalt-andesitic lava flow	49	0.033	1178348	0.063	0.535
	6. Colluvial slope of debris	0	0.000	31466	0.002	0.000
	7. Piedmont	10	0.007	265587	0.014	0.484
	8. Dissected pyroclastic flow surface	8	0.005	131593	0.007	0.782
	9. Floodplain	1	0.001	38939	0.002	0.330
	10. Glacial slope	0	0.000	573	0.000	0.000
	11. Glacier valley	0	0.000	34690	0.002	0.000
	12. Glacio-fluvial slope	27	0.018	1383588	0.073	0.251
	13. Hill and ridge dissected into sedimentary rock	76	0.052	484209	0.026	2.020
	14. Hill and ridge in intrusive rock	5	0.003	23591	0.001	2.727
	15. Hill and ridge in metamorphic rock	5	0.003	95887	0.005	0.671
	16. Hill and ridge in sedimentary rock	25	0.017	311836	0.017	1.032
	17. Hill and ridge in volcanic rock	30	0.020	1065009	0.057	0.362
	18. Hill and ridge in volcano-sedimentary rock	8	0.005	264565	0.014	0.389
	19. Hill in intrusive or subvolcanic body	2	0.001	20638	0.001	1.247
	20. Hill in intrusive rock	0	0.000	11115	0.001	0.000
	21. Hill in sedimentary rock	2	0.001	139272	0.007	0.185
	22. Hill in volcanic rock	2	0.001	36717	0.002	0.701
	23. Ignimbritic hill and ridge	0	0.000	55612	0.003	0.000

	24. Lagoon and bodies of water	0	0.000	58417	0.003	0.000
	25. Monogenetic slag cone	0	0.000	12977	0.001	0.000
	26. Moraines	3	0.002	716619	0.038	0.054
	27. Mountain in intrusive rock	110	0.075	678845	0.036	2.085
	28. Mountain in metamorphic rock	43	0.029	268824	0.014	2.058
	29. Mountain in sedimentary rock	49	0.033	444798	0.024	1.418
	30. Mountain in volcanic rock	104	0.071	1339979	0.071	0.999
	31. Mountain in volcano-sedimentary rock	8	0.005	48491	0.003	2.123
	32. Mountain slope in intrusive rock	34	0.023	111197	0.006	3.934
	33. Mountain with glacier cover	5	0.003	67755	0.004	0.950
	34. Mountains and hills in intrusive rock	12	0.008	106325	0.006	1.452
	35. Mountains and hills in metamorphic rock	8	0.005	40051	0.002	2.570
	36. Mountains and hills in sedimentary rock	22	0.015	179922	0.010	1.573
	37. Mountains and hills in volcanic rock	100	0.068	743496	0.039	1.731
	38. Piedmont fan	2	0.001	625924	0.033	0.041
	39. Pyroclastic flow surface	232	0.158	2411997	0.128	1.238
	40. Shed with slip tank	46	0.031	144804	0.008	4.088
	41. Alluvial-lacustrine foothills	0	0.000	14598	0.001	0.000
	42. Alluvial foothills	5	0.003	30470	0.002	2.112
	43. Stratovolcano	151	0.103	1535309	0.081	1.266
	44. Structural hill in sedimentary rock	5	0.003	65153	0.003	0.987

	45. Structural mountain in sedimentary rock	67	0.046	911974	0.048	0.945
	46. Structural mountains and hills in sedimentary rock	120	0.082	1265022	0.067	1.221
	47. Undifferentiated terrace	0	0.000	24672	0.001	0.000
	48. Volcanic caldera	0	0.000	12592	0.001	0.000
	49. Volcanic complex	19	0.013	321844	0.017	0.760
	50. Volcanic dome	25	0.017	276800	0.015	1.162
	51. Wetland	0	0.000	5226	0.000	0.000
Hydrogeology	1. Cambrian marine-sedimentary	0	0.000	45272	0.002	0.000
	2. Continental Cretaceous marine	30	0.020	378957	0.020	1.019
	3. Continental lower carboniferous	6	0.004	21030	0.001	3.671
	4. Continental Pleistocene quaternary	70	0.048	1214383	0.064	0.742
	5. Continental upper Cretaceous Paleogene	3	0.002	12979	0.001	2.974
	6. Continental upper Cretaceous	0	0.000	20446	0.001	0.000
	7. Cretaceous superunit Incahuasi	0	0.000	6619	0.000	0.000
	8. Cretaceous upper marine	9	0.006	418288	0.022	0.277
	9. Devonian marine	4	0.003	7132	0.000	7.217
	10. Holocene-continental quaternary	92	0.063	1690428	0.090	0.700
	11. Jurassic upper Cretaceous marine continental	114	0.078	1245882	0.066	1.177
	12. Lagoon	0	0.000	41367	0.002	0.000

	13. Middle Jurassic marine	3	0.002	61736	0.003	0.625
	14. Miocene-continental Neogene	158	0.108	1457107	0.077	1.395
	15. Miocene-volcanic Neogene	19	0.013	729772	0.039	0.335
	16. Miocene Neogene volcano-sedimentary	105	0.072	2232297	0.118	0.605
	17. Neogene Paleogene volcano-sedimentary	284	0.194	2989766	0.159	1.222
	18. Paleogene-Neogene, marine	16	0.011	44294	0.002	4.648
	19. Paleogene tonalite and granodiorite	9	0.006	43298	0.002	2.675
	20. Paleogeno-paleocene continental	23	0.016	224197	0.012	1.320
	21. Paleozoic granodiorite-granite	3	0.002	32816	0.002	1.176
	22. Pliocene-volcanic Neogene	50	0.034	540091	0.029	1.191
	23. Precambrian	0	0.000	16586	0.001	0.000
	24. Precambrian gneiss	125	0.085	819048	0.043	1.964
	25. Quaternary-volcanic Neogene	200	0.137	2942953	0.156	0.874
	26. Quaternary, Pliocene, Holocene, volcanic	48	0.033	771210	0.041	0.801
	27. Subvolcanic body	27	0.018	198889	0.011	1.747
	28. Superunit, Tiabaya	43	0.029	549465	0.029	1.007
	29. Upper Cretaceous monzogranite	23	0.016	81904	0.004	3.613
Landuse/landcover (LULC)	1. Water	0	0.000	107047	0.006	0.000
	2. Vegetation/Forest	2	0.001	39675	0.002	0.649
	3. Grassland	1	0.001	24594	0.001	0.523

	4. Flooded vegetation	0	0.000	8604	0.000	0.000
	5. Agricultural land	38	0.026	492148	0.026	0.994
	6. Scrubland	874	0.597	13535328	0.718	0.831
	7. Built-up area	8	0.005	137670	0.007	0.748
	8. Barren land	535	0.365	4457645	0.237	1.545
	9. Snow/ice cover	6	0.004	38192	0.002	2.022
	10. Cloud cover	0	0.000	200	0.000	0.000
Soil type	0. Acrisols	30	0.020	281178	0.015	1.373
	3. Andosols	2	0.001	41764	0.002	0.616
	4. Arenosols	10	0.007	128080	0.007	1.005
	5. Calcisols	11	0.008	233896	0.012	0.605
	6. Cambisols	681	0.465	9504227	0.504	0.922
	7. Chernozems	0	0.000	347	0.000	0.000
	11. Fluvisols	13	0.009	78787	0.004	2.123
	13. Gypsisols	0	0.000	1273	0.000	0.000
	15. Kastanozems	0	0.000	107	0.000	0.000
	16. Leptosols	315	0.215	4769218	0.253	0.850
	18. Luvisols	5	0.003	78476	0.004	0.820
	20. Phaeozems	0	0.000	4289	0.000	0.000
	24. Regosols	393	0.268	3482370	0.185	1.452
	25. Solonchaks	4	0.003	234257	0.012	0.220
	29. Vertisols	0	0.000	2290	0.000	0.000
Distance to faults (km)	1. <0.05	192	0.131	2287008	0.121	1.080
	2. 0.05 – 1	192	0.131	2090639	0.111	1.182
	3. 1 – 2	347	0.237	3435481	0.182	1.300
	4. 2 – 5	425	0.290	6047734	0.321	0.904
	5. 5 – 10	262	0.179	4185400	0.222	0.806
	6. 10 – 15	43	0.029	709258	0.038	0.780
	7. 15 – 20	3	0.002	82957	0.004	0.465
	8. 20 – 25	0	0.000	7	0.000	0.000
Earthquake magnitude	1. 3.40 – 4.28	229	0.156	3008355	0.160	0.980
	2. 4.28 – 4.48	620	0.423	7640159	0.406	1.044
	3. 4.48 – 4.72	415	0.283	5702655	0.303	0.936
	4. 4.72 – 5.07	168	0.115	2071640	0.110	1.044
	5. 5.07 – 6.90	32	0.022	417750	0.022	0.986
Distance to road	1. <0.05	417	0.285	3918466	0.208	1.369
	2. 0.05 – 2	501	0.342	6285881	0.334	1.026
	3. 2 – 5	347	0.237	5787463	0.307	0.772
	4. 5 – 10	126	0.086	2041601	0.108	0.794
	5. 10 – 20	73	0.050	756725	0.040	1.241

	6. 20 – 30	0	0.000	48348	0.003	0.000
Distance to epicenter (km)	1. <0.05	15	0.010	247534	0.013	0.780
	2. 0.05 – 1	77	0.053	715032	0.038	1.386
	3. 1 – 2	244	0.167	2558772	0.136	1.227
	4. 2 – 5	745	0.509	9395268	0.499	1.021
	5. 5 – 10	344	0.235	5191481	0.276	0.853
	6. 10 – 15	29	0.020	540542	0.029	0.690
	7. 15 – 20	9	0.006	161385	0.009	0.718
	8. 20 – 25	1	0.001	31734	0.002	0.406

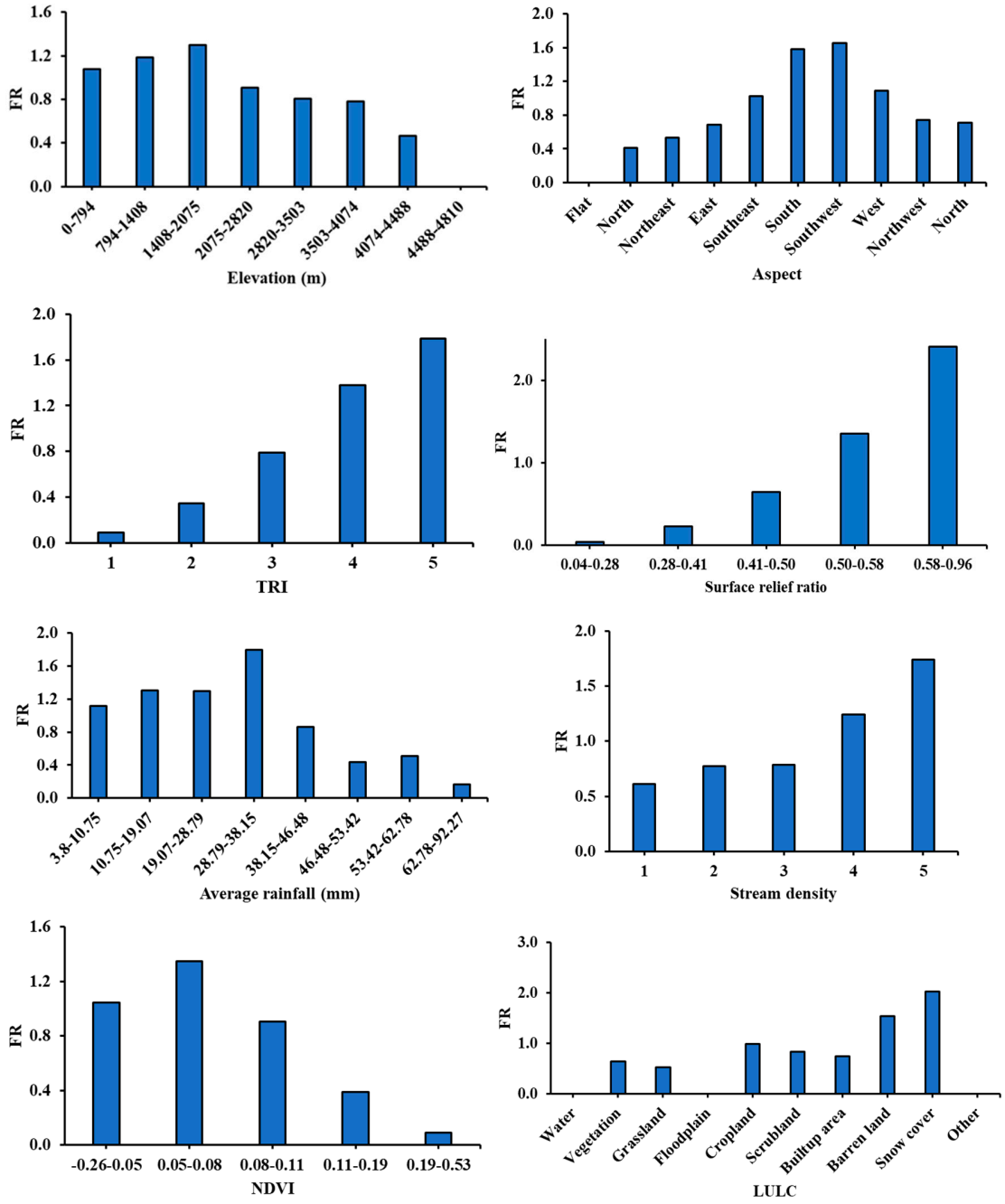


Figure S2. (continue on next page)

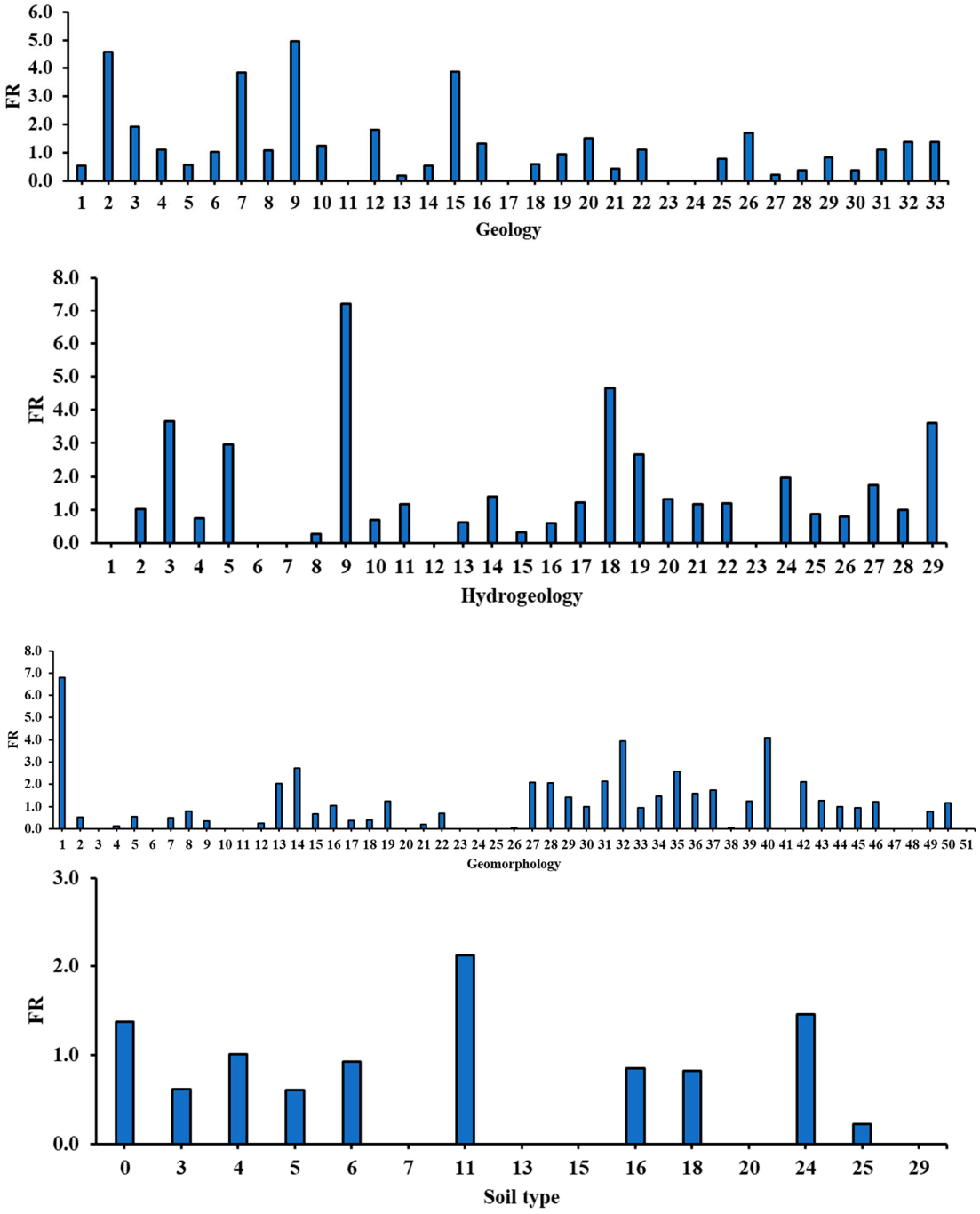


Figure S2. (continue on next page)

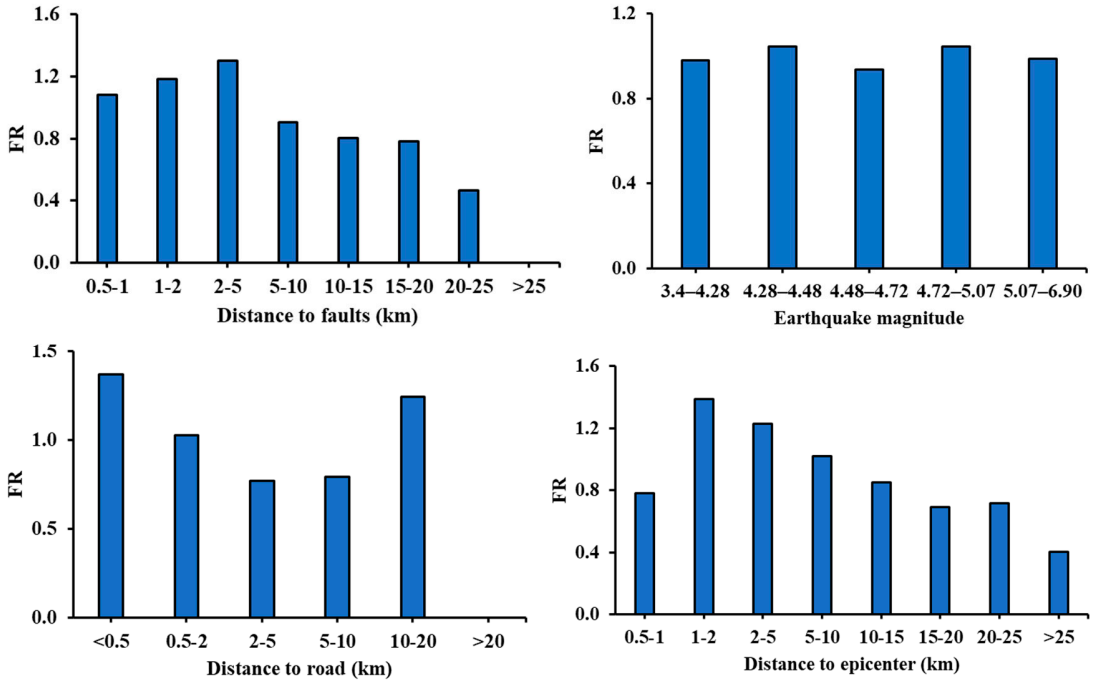


Figure S2. Frequency Ratio (FR) plots of different LIFs.