CrossMark

ORIGINAL ARTICLE

GIS-based landslide susceptibility assessment in Seoul, South Korea, applying the radius of influence to frequency ratio analysis

Jin Son[1] · Jangwon Suh[2] · Hyeong-Dong Park[1]

Abstract The objective of this paper is to map landslide susceptibility using a statistical analysis model and the radius of influence within a geographic information systems environment. The statistical analysis included triggering factors (e.g., topography, land cover, forest, and soil properties) of adjacent areas, in addition to the landslide sites themselves. To estimate the probability of landslide occurrence using the radius of influence, and to produce a landslide susceptibility index (LSI), we performed frequency radio (FR) analysis by applying the radius of influence to the domain of specific training sites. Landslide susceptibility maps were generated for each radius of influence, ranging from 0 to 300 m in 30 m increments. We observed enhanced FR index values corresponding to reduced exaggeration of statistical anomalies within the proper radius of influence. It is referred that by adopting the radius of influence the classes that not only affect the landslide occurrence from the adjacent areas but also make anomaly errors can be taken into account in FR analysis. Moreover, comparing the FR values between adopting the optimum radius of influence or not, we inferred that the greater the gap, the bigger influence of adjacent areas the classes have. In the validation stage, we identified the optimum radius of influence by measuring the area beneath the relative operating characteristics curve. We found that the optimum radius of influence in the study area is 240 m,

for which the LSI map is 5.95 % points more accurate than when not considering the radius of influence.

Keywords Landslide · GIS · Frequency ratio · Radius of influence

Introduction

Landslides are one of the principal geological disasters in urban areas, causing substantial property damage and human casualties. Urban areas are particularly vulnerable to landslides due to ground instability caused by human activity. Moreover, the high population density of urban areas means more people are exposed to geohazards than in rural regions. In recent years, landslide frequency has steadily increased due to intensifying rainfall, likely driven by ongoing climate change (Guzzetti et al. 2007; Kim et al. 2011, 2015). It is vital, therefore, to identify landslide-susceptible areas in order to mitigate and prepare for landslide events. In this regard, several studies have already made progress in landslide prediction based on geostatistical approaches.

The majority of research to date has used geographic information systems (GIS) to assess landslide susceptibility, as this is an effective approach to the quantitative evaluation of large regions. For example, landslide susceptibility has been estimated using GIS for several regions of South Korea where the risk of landslide hazards is particularly severe (Choi et al. 2012; Lee et al. 2006, 2013; Lee and Lee 2005; Lee and Min 2001; Park et al. 2013; Suh et al. 2011). Most of these investigations used GIS to perform spatial statistical analyses on susceptible areas. Other studies have employed statistical methods to evaluate landslide susceptibility; e.g., by comparing approaches

✉ Hyeong-Dong Park
hpark@snu.ac.kr

1 Department of Energy Systems Engineering, College of Engineering, Seoul National University, Seoul, South Korea

2 Department of Energy and Mineral Engineering, College of Earth and Mineral Sciences, The Pennsylvania State University, University Park, PA, USA

including frequency ratio (FR) analysis, analytic hierarchical processes (AHP), artificial neural networks (ANN), and logistic regression (LR). Additionally, principal triggering factors of past landslides have been clarified and verified by comparison with a landslide inventory database (Akgun and Türk 2010; Oh and Pradhan 2011; Park et al. 2013; Pouduyal et al. 2010; Pradhan 2011; Reis et al. 2012; Schleier et al. 2014; Shahabi et al. 2015; Suh et al. 2011; Yalcin et al. 2011; Yilmaz 2009, 2010; Yilmaz et al. 2012).

Most previous studies have concentrated on assessing susceptible areas by analyzing their geographical and geological properties, especially focused on identifying the spatial pattern of landslides without considering the properties of adjacent areas (Goltz 1996; Liucci et al. 2015; Malamud et al. 2004). Yet the physical characteristics of the surrounding area can influence landslide hazards in both direct and indirect ways (Jones and Preston 2012; Massey et al. 2013; Petley et al. 2005; Van Westen et al. 2003). Here, we present a statistical approach for evaluating landslide susceptibility that considers triggering factors both at the landslide site and in the surrounding area, and that enables the production of landslide susceptibility index (LSI) maps with varying radii of influence. To find the optimum radius of influence, we used the area under the relative operating characteristics curve (AUC) method to compare LSI maps with selected control areas of landslide inventories.

The objectives of this paper are to (1) demonstrate our evaluation model of landslide susceptibility, which utilizes GIS technology and statistics; (2) estimate the LSI by incorporating the effects of adjacent areas and varying the radius of influence; and (3) identify the optimum radius of influence and to validate this approach by comparing our results with those of previous studies.

Study area

The study area is the Seoul metropolitan area (605.2 km^2), which, with 10.6 million inhabitants, is the most urbanized area in South Korea. Seoul is located between the 37°25′N and 37°42′N, 126°45′E and 127°10′E, and is surrounded by mountainous terrain. Together with regular typhoons and intense summer rainfall, this high-relief environment results in a high incidence of landslides. For example, 69 landslides were recorded following intense rainfall (maximum hourly rainfall of 110.5 mm) on 27 July 2011, with more than half of these occurring close to Mt. Umyeon in the central-southern mountainous region of Seoul.

Geological structure of Mt. Umyeon, Seoul is mostly composed of gneiss and granite covered by soil layer, which has relatively stable bedrock. Yet infrastructures made by human activities have expanded to the steep

mountainous area, slope failures including rock falls and debris flows occur in rainy season. Landslides in Mt. Umyeon in 2011 were triggered by heavy rainfall which deteriorated the surface stability that result in soil failure. In the same period of time, other landslide spots were in similar situation, most damages were made by soil failure. We assumed that the amount of precipitation over the study area was almost equal, so that the triggering factors, such as topography, soil, forest properties and land cover, which are known as highly involved with soil failure including debris flow, were chosen. Accordingly, the landslide susceptibility assessment of this study is focused on the case of landslides triggered with same precipitation condition, especially in urban areas on mountainous terrain.

A landslide inventory database is maintained by the Seoul metropolitan government and includes location and scale data for each landslide event. Location data include text address and approximate longitude/latitude information. The spatial extent of each event is recorded in the scale field, though the coverage is not precise. To specify the spatial coverage, the exact locations of landslide sites are assigned by manual geocoding and coverage is represented by a circular area, the diameter of which is calculated from the spatial scale of the event. As the landslide inventory is presented in map form, data are validated against seven known landslide sites, which constitute approximately 10 % of all landslide events (Fig. 1). The validation sites are sampled from a uniform distribution that properly represents each cluster. The remaining 62 landslide sites are used as training data for statistical analysis.

Building the triggering factor database

In this study, 12 geospatial attributes are considered to be triggering factors on the basis of their geological and geographical properties (Table 1): slope angle, slope aspect, curvature, land cover, forest age, forest density, forest diameter, forest type, soil drainage, soil material, soil texture, and soil thickness. In accordance with Lee and Lee (2005), we selected these triggering factors for the following reasons: (1) spatial data on triggering factors are obtainable from government organizations; (2) the study areas are sufficiently close to one another, meaning that the same triggering factors are relevant at both; and (3) the LSI assessment method can be assessed effectively since we have employed the same statistical processes. Spatial data, including topographical maps, geological maps, and forest/soil properties maps, have been analyzed to rank the landslide prone areas.

Entries of database are selected as major factors that reflect the characteristics of landslide events in South

Fig. 1 Locations of landslide sites (for training and validation) in the study area
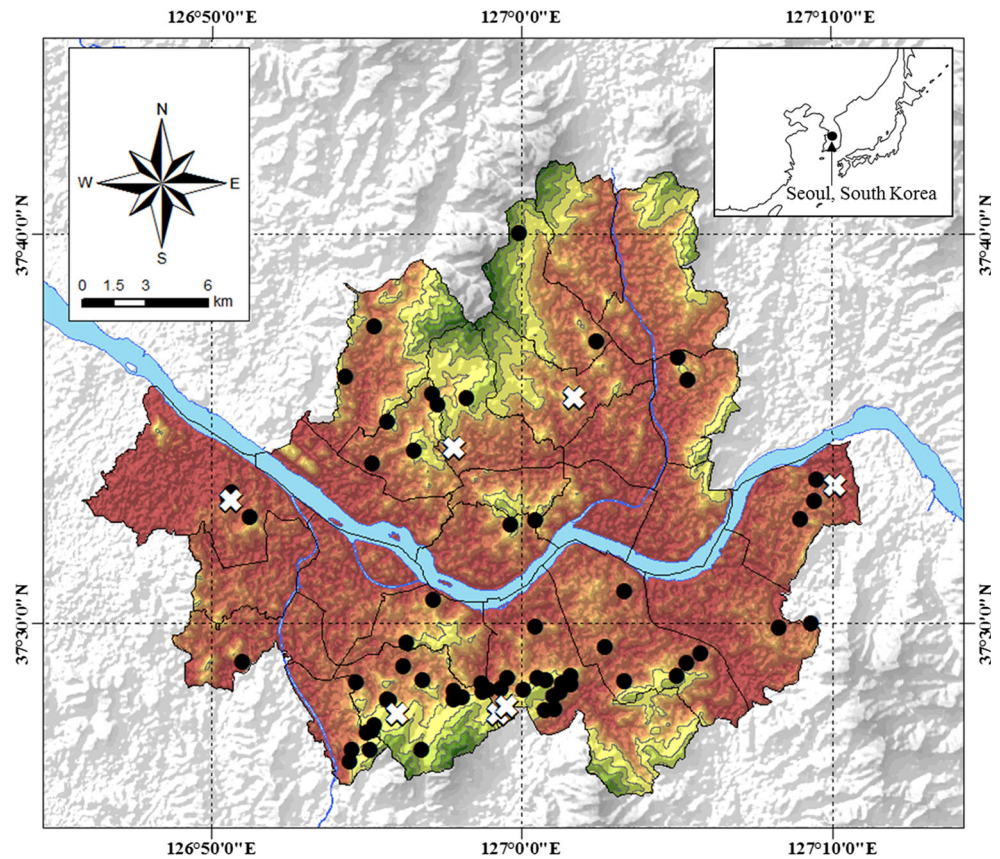
Table 1 Landslide inventory and geospatial attributes database of the study area

Classification	Factor	Data type	Scale	Distributed by
Landslide inventory	Location of landslide sites	Point	Area in m^2 without coverage	Seoul Metropolitan Government
	Training sites	Polygon	–	
	Validation sites	Polygon	–	
Topographic map	Slope angle	Grid	30 m × 30 m	ASTER G-DEM
	Slope aspect			
	Curvature			
Land cover map	Land cover	Polygon	–	WAMIS
Forest properties map	Forest age	Grid	30 m × 30 m	
	Forest density			
	Forest diameter			
	Forest type			
Soil properties map	Soil material	Grid	30 m × 30 m	
	Soil texture			
	Soil thickness			

Korea. Generally, slope angle, slope aspect and curvature are topographic attributes that are commonly suggested as landslide triggering factors (Schleier et al. 2014; Shahabi et al. 2015; Van Westen et al. 2003). The distance from the fault is also an important factor but there was no spatial database to be excluded. Seoul, the study area, has numerous soil failure, including debris flow and debris slide to cause the most dominant damages, land cover, forest and soil properties are chosen as triggering factors which have relation with inducing soil failure. Especially, previous studies in South Korea (Choi et al. 2012; Lee and Lee 2005; Lee et al. 2006; Yune et al. 2013), forest and soil properties are found as effective to accelerate or prevent from landslide occurrences.

To avoid the redundancy of the dataset, the classes which contain similar properties are required to be checked before FR analysis. In the entries of constructed database, forest age and forest diameter seem to be in close correlation but according to forest type, they could not be interdependent to each other. Likewise both soil texture and soil material are a similar kind of lithology attributes but they could show different aspects in surface drainage or moisture absorptivity depending on the combination. Hence, we conduct FR analysis to define them whether independent from landslide occurrence or not, and design more sophisticated analysis model.

All topographical data are constrained using a digital elevation model (DEM; ASTER G-DEM V2) with a resolution of ~30 m. This globally constant DEM is freely available from the Japanese Ministry of Economy, Trade and Industry (METI) and the US National Aeronautics and Space Administration (NASA). Maps of forest/soil properties and land cover are available from the Korean Water Management Information System (WAMIS; http://www.wamis.go.kr/) in GRID format and can be converted to raster files with a spatial resolution of 30 m.

The values of triggering factors are regarded as being either discrete or continuous. Basically, the classes of each triggering factor are classified with discrete (categorical) values, such as land cover and forest/soil properties. We tried to describe the whole classes in quantitative values even constructed with discrete values. Though, soil drainage was the only factor that cannot be numerated due to the surveys were made by naked eye of investigators who have specialized in over 20 years (Hong et al. 2009). Except this case, the quantitative factors, which are investigated with certain categories as constant interval method (i.e. forest age, forest density, forest diameter, and soil thickness), are used as the classes of the factors in FR analysis.

For the trigger factors comprising continuous values but generally defined with discrete classes, each factor map is reclassified into several classes using a constant interval method: flat and eight directions for aspect, and flat,
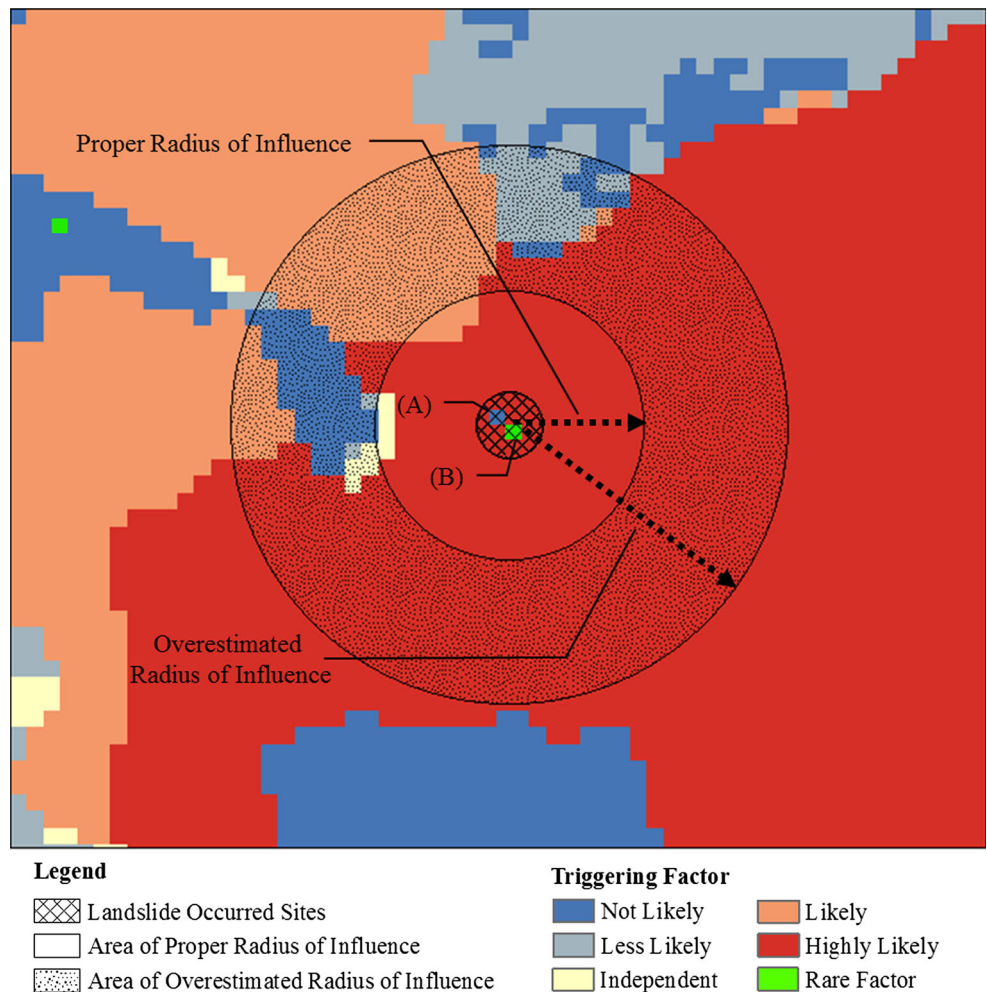
concave, and convex for curvature. However, in the case of slope angle data, use of the constant interval method results in the data of each class being strongly skewed. To equalize the irregular distribution, it is common method in classification to apply natural breaks (Jenks optimization), quantile interval or standard deviation. Comparing the results from adopting these methods, the standard deviation method is used for the reclassification and we manually set the number of domains to be similar in each class. In the occasion of the data are distributed irregularly, natural breaks or equal-constant interval methods cannot classify the number of domain to each class properly. On the other hand, equal-area interval method can solve the unbalanced distribution but suggest ambiguous border that hardly show the distinct characteristics of classes as excessively focusing on equality of the domain. If certain class has not sufficient samples or the number of classes of certain triggering factor is not enough, it can be supplemented by applying other methods including natural breaks, equal-constant interval or equal-area interval method.

## Concept of the radius of influence

Our approach differs from previous assessments of landslide susceptibility in that we incorporate the radius of influence into our GIS-based statistical analysis. Because of the likelihood that geohazards such as landslides can be caused by triggering factors in adjacent areas (Massey et al. 2013; Petley et al. 2005), it is more effective and accurate for landslide susceptibility training domains to cover an area extending beyond the immediate site of the event (Fig. 2). Adopting this concept is particularly advantageous in the case of landslides occurring at sites where triggering factors are deemed "less likely" or "rare," such as at points (A) and (B), respectively, in Fig. 2. Triggering factors can be exaggerated in probability calculations due to a lack of sufficient domains, resulting in erroneous assessments of landslide vulnerability and reduced accuracy of statistical analysis. On the other hand, the area within an overestimated radius of influence might contain factors that do not contribute to landslide occurrence. Such exaggerations (i.e., if the radius of influence is not applied or is overestimated) can be reduced.

In summary, it is crucial to find the optimum radius of influence in order to produce the most accurate results. This can be determined in the validation stage, following statistical analysis and production of the LSI maps. In the present study, we calculated different radii of influence at an interval of 30 m, which is the minimum spatial resolution of available property maps. To determine the optimum radius of influence for triggering factors, we conducted statistical analyses over 11 different radii, from 0 to 300 m (Fig. 3).

Fig. 2 Conceptual diagram showing the effects of applying the radius of influence to a statistical analysis of landslide susceptibility

Legend

⊠ Landslide Occurred Sites
☐ Area of Proper Radius of Influence
⣿ Area of Overestimated Radius of Influence

Triggering Factor

■ Not Likely   ■ Likely
■ Less Likely   ■ Highly Likely
☐ Independent   ■ Rare Factor

Fig. 3 Spatial database and process model for assessing landslide susceptibility in this study
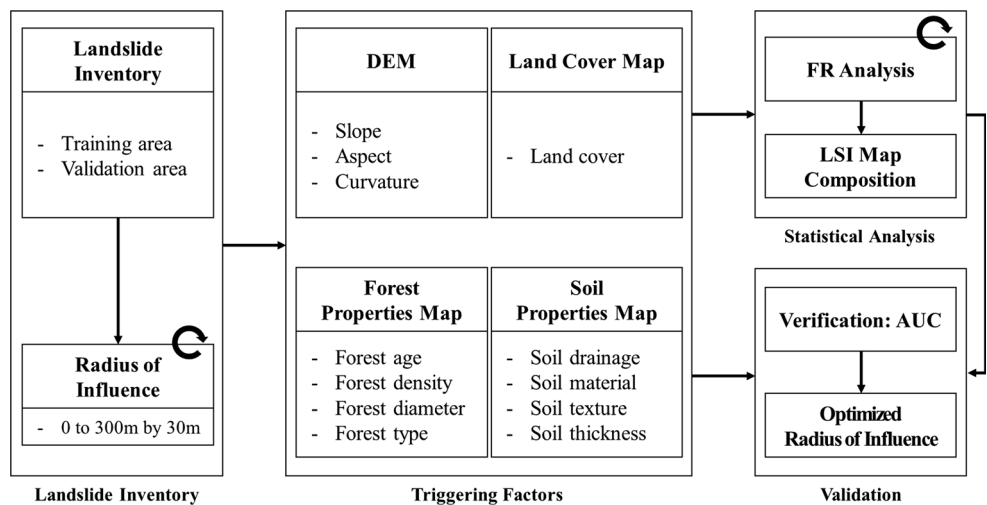
Frequency ratio analysis for each area of influence

To quantify landslide susceptibility, we used FR analysis. This statistical approach is commonly applied to susceptibility assessments in cases when both the landslide inventory database and triggering factor maps are available for the entire study area (Choi et al. 2012; Park et al. 2013; Reis et al. 2012; Schleier et al. 2014; Shahabi et al. 2015; Suh et al. 2011; Ulrich et al. 2008; Yalcin et al. 2011; Yilmaz 2009). Specifically, the FR method estimates the

Table 2 FR index values for each class of triggering factor according to the radius of influence

Factor	Class	Radius of influence										
		0 m	30 m	60 m	90 m	120 m	150 m	180 m	210 m	240 m	270 m	300 m
Slope	0.00°–2.01°	0.041	0.066	0.105	0.126	0.154	0.168	0.184	0.207	0.230	0.249	0.261
	2.01°–5.05°	0.186	0.297	0.354	0.401	0.426	0.449	0.475	0.497	0.516	0.543	0.569
	5.05°–8.08°	0.779	0.838	0.921	0.958	0.986	0.994	1.007	1.015	1.039	1.047	1.053
	8.08°–11.1°	1.499	1.583	1.522	1.492	1.451	1.491	1.503	1.511	1.499	1.478	1.482
	11.1°–14.1°	2.054	1.901	1.815	1.809	1.880	1.879	1.823	1.778	1.757	1.743	1.718
	14.1°–17.2°	2.773	2.397	2.312	2.216	2.105	2.114	2.135	2.105	2.048	1.996	1.941
	17.2°–20.2°	3.390	2.670	2.381	2.274	2.204	2.055	1.921	1.866	1.826	1.786	1.767
	20.2°–56.0°	2.414	2.373	2.163	2.004	1.869	1.739	1.672	1.616	1.548	1.493	1.446
Aspect	Flat	0.082	0.070	0.117	0.188	0.244	0.273	0.286	0.326	0.356	0.373	0.396
	N	0.698	0.840	0.906	0.893	0.895	0.900	0.941	0.974	1.009	1.033	1.060
	NE	0.909	0.871	0.831	0.805	0.818	0.824	0.837	0.849	0.864	0.881	0.881
	E	1.289	1.259	1.163	1.066	0.988	0.926	0.886	0.856	0.827	0.813	0.800
	SE	1.055	0.913	0.843	0.793	0.753	0.726	0.693	0.665	0.645	0.640	0.647
	S	0.397	0.441	0.515	0.576	0.604	0.614	0.623	0.645	0.660	0.677	0.691
	SW	0.272	0.474	0.560	0.642	0.677	0.705	0.740	0.770	0.785	0.793	0.807
	W	0.881	0.815	0.786	0.806	0.868	0.926	0.938	0.959	0.992	1.014	1.022
	NW	1.327	1.302	1.397	1.494	1.521	1.550	1.557	1.530	1.500	1.468	1.445
Curvature	Concave (−)	1.927	1.656	1.556	1.488	1.472	1.423	1.367	1.336	1.316	1.294	1.292
	Flat	0.659	0.708	0.718	0.737	0.742	0.750	0.760	0.769	0.778	0.790	0.800
	Convex (+)	1.271	1.186	1.287	1.269	1.319	1.355	1.383	1.391	1.395	1.366	1.342
Land cover	Water	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Urban area	0.148	0.218	0.275	0.337	0.384	0.431	0.473	0.506	0.535	0.558	0.578
	Bare soil	0.198	0.201	0.205	0.263	0.335	0.355	0.396	0.460	0.498	0.525	0.555
	Wet land	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Grass land	0.582	0.557	0.584	0.565	0.510	0.525	0.519	0.513	0.516	0.525	0.537
	Mountainous terrain	2.691	2.513	2.395	2.302	2.232	2.146	2.077	2.022	1.975	1.937	1.897
	Rice field	0.818	1.014	1.018	0.936	0.920	0.897	0.859	0.853	0.851	0.850	0.853
	Farm land	0.972	1.111	1.242	1.200	1.137	1.143	1.133	1.112	1.089	1.102	1.128
Forest age (>50 % of whole timber)	0–10 years	19.64	13.10	9.531	7.656	5.889	4.639	3.688	3.079	2.629	2.257	1.971
	10–20 years	1.839	1.993	2.067	2.149	2.289	2.360	2.364	2.364	2.384	2.412	2.434
	20–30 years	3.622	3.352	3.138	2.955	2.812	2.698	2.605	2.507	2.413	2.328	2.250
	30–40 years	0.292	1.039	1.957	2.115	1.966	1.892	1.878	1.807	1.651	1.498	1.364
	40–50 years	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016
Forest density	<50.0 %	0.603	0.919	1.163	1.294	1.411	1.650	1.769	1.879	2.066	2.111	2.155
	50.0–75.0 %	1.911	1.962	1.952	1.935	1.901	1.874	1.854	1.816	1.774	1.738	1.705
	>75.0 %	9.018	8.104	7.652	7.155	6.824	6.498	6.143	5.824	5.543	5.297	5.078
Forest diameter	<6 cm	19.52	12.93	9.549	7.619	5.931	4.728	3.806	3.172	2.737	2.377	2.073

Table 2 continued

Factor	Class	Radius of influence										
		0 m	30 m	60 m	90 m	120 m	150 m	180 m	210 m	240 m	270 m	300 m
	6–16 cm	3.162	3.047	2.932	2.832	2.753	2.679	2.609	2.531	2.453	2.385	2.324
	16–28 cm	0.143	0.583	1.005	1.039	0.986	0.961	0.947	0.888	0.828	0.750	0.712
Forest type	Larix Kaempferi	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Mixed forest	0.306	0.681	0.988	1.115	1.189	1.267	1.277	1.266	1.240	1.223	1.219
	Pinus Densiflora	2.990	3.453	3.559	3.391	3.192	3.053	2.922	2.709	2.484	2.276	2.109
	Hardwood forest	5.139	4.641	4.257	3.960	3.732	3.549	3.407	3.262	3.125	2.998	2.876
	Farm land	0.113	0.166	0.196	0.221	0.242	0.261	0.282	0.307	0.333	0.355	0.374
	Pinus Rigida	0.811	1.149	1.429	1.691	1.867	1.903	1.957	2.011	2.118	2.244	2.352
	Pinus Koraiensis	20.76	13.66	9.854	7.616	5.858	4.616	3.720	3.105	2.650	2.275	2.037
	Non-forest land	0.000	0.000	0.000	0.043	0.079	0.106	0.153	0.200	0.236	0.271	0.297
	Artificial hardwood forest	15.17	14.40	13.56	14.25	15.65	15.62	15.07	15.23	15.43	15.28	15.02
	Populus Deltoides	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Soil drainage (observation made by specialists' naked eye)	Excessively	2.209	2.175	2.182	2.183	2.181	2.162	2.128	2.093	2.056	2.018	1.983
	Well	8.233	8.724	8.918	8.987	8.777	8.540	8.231	7.976	7.812	7.621	7.410
	Moderately	0.437	0.469	0.478	0.478	0.474	0.481	0.490	0.505	0.519	0.537	0.559
	Rarely	0.000	0.007	0.021	0.044	0.091	0.124	0.155	0.170	0.179	0.196	0.207
	Poorly	0.186	0.194	0.202	0.219	0.246	0.274	0.310	0.346	0.378	0.409	0.431
	Outcrop	0.540	0.576	0.567	0.578	0.578	0.581	0.605	0.622	0.643	0.652	0.666
Soil material	Alluvium	0.024	0.036	0.047	0.057	0.078	0.088	0.108	0.120	0.132	0.145	0.157
	Valley alluvium	0.000	0.023	0.035	0.046	0.056	0.068	0.078	0.085	0.092	0.098	0.103
	Gley alluvium	0.189	0.194	0.200	0.217	0.241	0.271	0.306	0.344	0.377	0.409	0.431
	Sedimentary soil	0.000	0.111	0.332	0.482	0.624	0.753	0.885	0.973	1.008	1.118	1.236
	Red yellow soil	0.055	0.195	0.309	0.365	0.393	0.415	0.411	0.408	0.411	0.427	0.432
	Lithosol and red yellow soil	0.712	0.712	0.683	0.657	0.637	0.631	0.636	0.650	0.663	0.682	0.709
	Lithosol	2.490	2.481	2.496	2.501	2.489	2.460	2.414	2.368	2.325	2.280	2.237
	Red yellow sedimentary soil	0.738	1.018	1.316	1.549	1.782	2.072	2.266	2.420	2.535	2.602	2.714
	Outcrop	0.540	0.576	0.567	0.578	0.578	0.581	0.605	0.622	0.643	0.652	0.666
Soil texture	Coarse Sand	0.012	0.030	0.041	0.047	0.057	0.065	0.077	0.086	0.096	0.107	0.115
	Silt loam	2.122	2.157	1.662	1.800	2.063	2.260	2.502	2.647	2.771	2.811	2.827
	Gravelly loam	0.000	0.000	0.000	0.117	0.306	0.618	0.944	1.240	1.429	1.479	1.453
	Clay sand	0.000	0.111	0.332	0.482	0.624	0.753	0.885	0.973	1.008	1.118	1.236
	Clay loam	0.229	0.213	0.212	0.198	0.195	0.195	0.196	0.209	0.228	0.251	0.271
	Clay	0.023	0.098	0.174	0.238	0.287	0.334	0.384	0.429	0.464	0.505	0.534
	Gravelly sand	0.000	0.000	0.000	1.049	2.239	2.955	3.639	3.672	3.423	3.279	3.207
	Sandy silt loam	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Coarse loam	2.087	1.984	1.933	1.896	1.861	1.834	1.803	1.782	1.764	1.750	1.738
	Rocky loam	0.738	1.018	1.316	1.549	1.782	2.072	2.266	2.420	2.535	2.602	2.714

Table 2 continued

Factor	Class	Radius of influence										
		0 m	30 m	60 m	90 m	120 m	150 m	180 m	210 m	240 m	270 m	300 m
	Sandy loam	0.797	1.122	1.280	1.375	1.427	1.430	1.420	1.393	1.363	1.323	1.296
	Outcrop	0.540	0.576	0.567	0.578	0.578	0.581	0.605	0.622	0.643	0.652	0.666
Soil thickness	>150 cm	0.059	0.208	0.330	0.371	0.344	0.332	0.307	0.308	0.325	0.349	0.367
	125–150 cm	0.228	0.212	0.211	0.200	0.206	0.212	0.217	0.228	0.245	0.267	0.285
	100–125 cm	0.512	0.512	0.491	0.475	0.464	0.462	0.469	0.483	0.496	0.514	0.536
	75–100 cm	2.122	2.157	1.662	1.800	2.063	2.260	2.502	2.647	2.771	2.811	2.827
	50–75 cm	0.068	0.128	0.199	0.272	0.347	0.429	0.518	0.593	0.645	0.690	0.724
	20–50 cm	1.897	1.896	1.910	1.916	1.910	1.890	1.857	1.824	1.793	1.760	1.728
	<20 cm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Outcrop	0.540	0.576	0.567	0.578	0.578	0.581	0.605	0.622	0.643	0.652	0.666

probability of each class of triggering factors through spatial comparison with landslide inventories, both at the landslide site and elsewhere. Ultimately, the FR index indicates the degree to which each class of triggering factors correlates with landslide occurrence. In the case that the landslide training domain is much smaller than the whole domain, the class of the triggering factor supports ground stability. If the situation is reversed, the FR index exceeds 1, indicating a vulnerability to landslide failure. If the FR index is close to 1, the class of triggering factor has no effect on the potential for landslide events.

The FR index can be determined using Eq. 1, in which each variable is assessed by overlaying the maps of spatial databases. For the sake of calculation, the number of grid cells in the landslide inventories and for the entire study area need to be measured. It is also necessary to count the grid cells for different classes of triggering factor that occur within each domain, both in the landslide inventory and over the whole study area. The FR index of each class of triggering factor can then be evaluated by inserting these values into the equation.

$$\text{Frequency Ratio(FR)} = \frac{\text{Percent of Landslide Training Domain for each class}}{\text{Percent of Whole Domain for each class}} \quad (1)$$

Accordingly, FR index layer maps of 12 factors are generated in 11 separate training domains for each radius of influence. The variables for each area of influence are measured by cross-tabbing training domains with maps of geological and geographical properties. While this process is time consuming when done manually, our spatial process module is designed to compute these analyses with multiple iterations of the radii of influence, resulting in FR index values for each class of triggering factors (Table 2).

Without considering the radius of influence, the most landslide-prone classes include mountainous terrain in the land cover category, <10 years old in forest age, <6 cm in forest diameter (i.e., the median trunk diameter at breast height), *Pinus koraiensis* and artificial hardwood forest in forest type, and well drained in soil drainage. Together, these characteristics suggest that landslides occur primarily in mountainous regions with a cover of young, relatively small trees that are unlikely to support ground stability.

As the radius of influence increases, several FR indices show a gradual decrease or increase, ultimately approaching a value of 1. Therefore, any previous exaggeration of these classes is reduced when the area of influence is taken into account. For example, for classes with an initial FR index value of 0, the value increases dramatically once the radius of influence expands. This explains why larger areas of influence yield larger sample numbers from the training sites. This approach can help reduce the exaggeration of data and improve the accuracy of LSI maps.

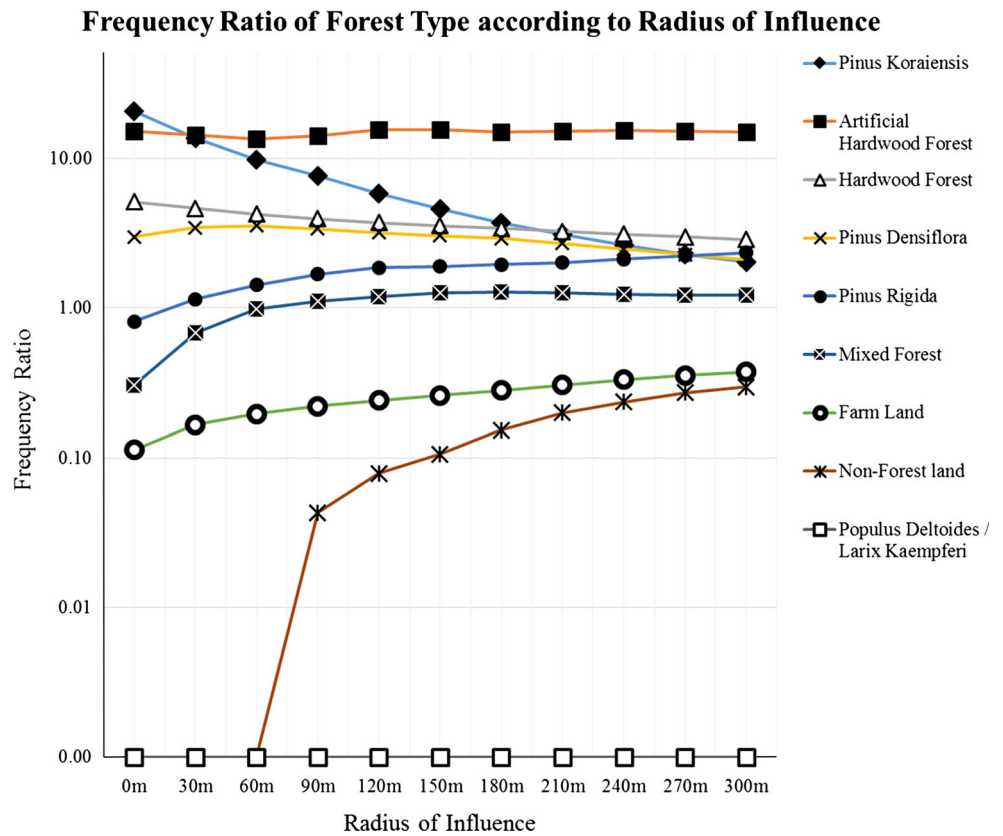**Fig. 4** Frequency ratio index values of different forest types according to the radius of influence

Frequency Ratio of Forest Type according to Radius of Influence

Figure 4 shows representative trends in FR index values for each class of forest type. The FR index values for Pinus koraiensis and artificial hardwood forests exceed 15 within a 0 m radius. However, the values for Pinus koraiensis drop sharply from 20.76 to 2.037 with increasing radius size, indicating that this forest type is not a principal triggering factor for landslides. In contrast, the FR index value of artificial hardwood forest does not decrease but remains close to 15 for every radius of influence, suggesting that this class contributes to landslides in a way that is largely independent of the radius of influence. Similarly, there are triggering factors that tend not to induce a landslide. For example, both non-forested land and Populus deltoides classes give values of zero within a 0 m radius. However, as the radius increases, the FR index value of non-forested land increases gradually from 0 to 0.297, which is not apparent for radii of <60 m. Meanwhile, the *Populus deltoides* class shows no change in FR with increasing radius, indicating that this factor is independent of the radius of influence.

Compilation of LSI maps

To calculate LSI values, the corresponding FR index values of the factors listed in Table 2 are calculated for each grid cell according to Eq. 2. Subsequently, LSI maps for each

radius of influence are generated by stretching the summed values of each grid cell from 1 to 100 in equal area. As shown in Fig. 5, mountainous regions in southern areas, where major landslides have occurred, are classed as highly susceptible. The southeastern and northern parts of the study area are classified as highly vulnerable, though landslides have not yet occurred in these regions. As the radius of influence increases, the LSI map becomes blurred, indicating that highly variable values are being smoothed. As depicted in Table 2, this smoothing effect occurs when the LSI value of each grid cell is deduced on the basis of compensated FR index values.

$$LSI = \sum FR \text{ of each factor for each radius of influence.} \quad (2)$$

Validation of the estimated optimal radius of influence

As landslide susceptibility maps for each radius of influence are useful predictive tools, it is important to evaluate which LSI map is the most accurate by comparison with validation data. The most common approach for making this comparison in geohazard contexts employs the relative operating characteristics (ROC) curve method and the area under the ROC curve (AUC) method (e.g., Choi et al. 2012; Lee et al. 2013; Mason and Graham 2002; Oh and
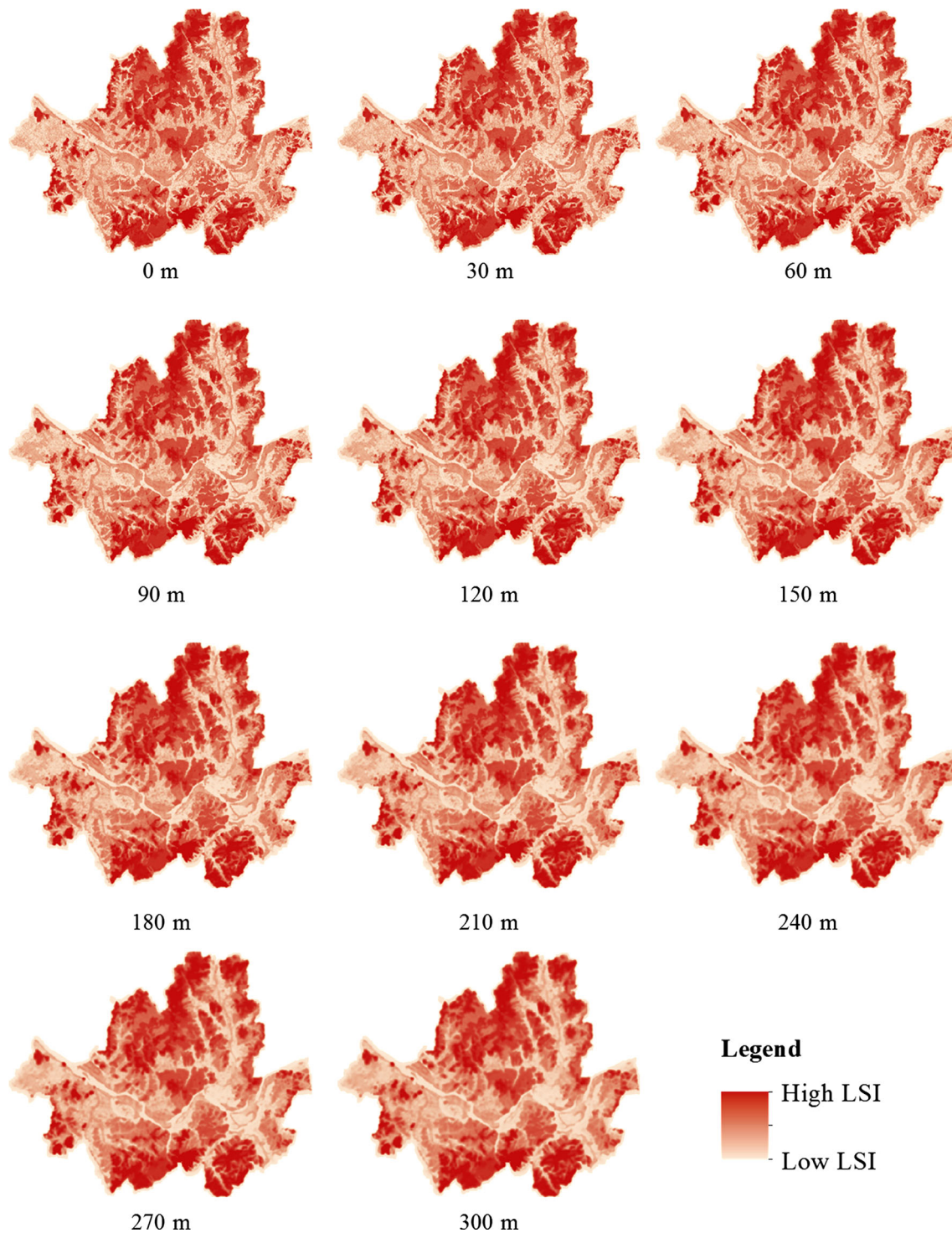
Fig. 5 Generated LSI maps for each radius of influence

Pradhan 2011; Suh et al. 2013; Yilmaz 2009). To draw the ROC curve, LSI maps for each radius of influence are combined with the landslide validation area. The cumulative frequency histogram is made using the ratio of corresponding grid cells to LSI, as shown in Table 3. To ensure

the AUC value shows a positive correlation with the accuracy of the LSI map, LSI is ranked in descending order. With increasing measured value of AUC, the average LSI of the validation sites also increases, which indicates the LSI map has greater accuracy.

Table 3 Cumulative frequency of LSI and accuracy of prediction maps according to the radius of influence

LSI in rank (%)	Radius of influence										
	0 m (%)	30 m (%)	60 m (%)	90 m (%)	120 m (%)	150 m (%)	180 m (%)	210 m (%)	240 m (%)	270 m (%)	300 m (%)
1	2.44	3.66	8.54	6.10	10.39	11.11	4.92	4.76	5.26	1.72	1.96
2	7.32	8.54	15.85	18.29	23.38	11.11	8.20	7.94	12.28	6.90	7.84
3	8.54	9.76	23.17	24.39	24.68	13.89	13.11	12.70	15.79	13.79	15.69
4	9.76	10.98	26.83	25.61	25.97	20.83	16.39	14.29	19.30	15.52	21.57
5	9.76	19.51	35.37	29.27	29.87	27.78	24.59	19.05	22.81	22.41	25.49
6	10.98	21.95	40.24	37.80	36.36	30.56	26.23	25.40	28.07	29.31	31.37
7	28.05	37.80	47.56	41.46	42.86	34.72	32.79	31.75	36.84	34.48	43.14
8	39.02	47.56	47.56	47.56	48.05	44.44	45.90	39.68	45.61	41.38	49.02
9	47.56	53.66	50.00	48.78	50.65	51.39	52.46	52.38	54.39	53.45	56.86
10	54.88	54.88	53.66	51.22	55.84	56.94	62.30	61.90	68.42	65.52	68.63
11	59.76	62.20	54.88	54.88	62.34	65.28	70.49	68.25	78.95	74.14	74.51
12	62.20	62.20	58.54	57.32	66.23	66.67	77.05	69.84	80.70	82.76	80.39
13	62.20	62.20	60.98	57.32	71.43	69.44	81.97	82.54	92.98	93.10	86.27
14	62.20	62.20	65.85	65.85	77.92	75.00	88.52	90.48	94.74	96.55	90.20
15	62.20	62.20	68.29	68.29	80.52	83.33	88.52	90.48	94.74	96.55	96.08
16	62.20	62.20	70.73	69.51	81.82	83.33	91.80	92.06	98.25	96.55	98.04
17	62.20	62.20	73.17	70.73	81.82	88.89	95.08	98.41	98.25	98.28	98.04
18	63.41	63.41	74.39	76.83	87.01	90.28	98.36	98.41	98.25	100.00	98.04
19	63.41	63.41	76.83	79.27	89.61	90.28	98.36	100.00	98.25		98.04
20	67.07	65.85	79.27	82.93	93.51	95.83	98.36		100.00		98.04
21	69.51	69.51	84.15	84.15	93.51	97.22	100.00				100.00
22	71.95	73.17	85.37	86.59	93.51	97.22					
23	74.39	74.39	85.37	86.59	94.81	98.61					
24	74.39	74.39	85.37	86.59	96.10	100.00					
25	84.15	75.61	86.59	86.59	97.40						
26	87.80	86.59	86.59	87.80	97.40						
27	89.02	90.24	89.02	90.24	98.70						
28	90.24	90.24	91.46	90.24	100.00						
29	90.24	90.24	91.46	93.90							
30	92.68	90.24	91.46	95.12							
31	92.68	91.46	91.46	96.34							
32	93.90	92.68	97.56	100.00							
33	96.34	97.56	100.00								
34	96.34	100.00									
35	100.00										
AUC	0.8649	0.8693	0.8898	0.8898	0.9112	0.9104	0.9175	0.9160	0.9244	0.9222	0.9239

As the radius of influence increases, both the AUC values and LSI values (e.g., minimum LSI, LSI of the top 80 %) also increase (Fig. 6). However, the magnitude of the increase declines considerably and AUC values fluctuate beyond a radius of 210 m. We speculate that this pattern indicates that triggering factors in adjacent areas have a measurable impact on landslide occurrence up to the radius of 210 m. Conversely, AUC values decrease slightly beyond 270 m, which we interpret as overestimation of the area of influence, resulting in the inclusion of irrelevant triggering factors in the analysis.

Through this verification process, we determined that the optimum radius of influence is 240 m, where the AUC value peaks. The AUC value of 86.49 % obtained using existing methods shows an increase to 92.44 % when considering the optimized radius of influence. Moreover, in the case of 0 m, the lowest LSI is 35 %, while for 240 m it is only 20 %. According to Yilmaz (2009), the lowest
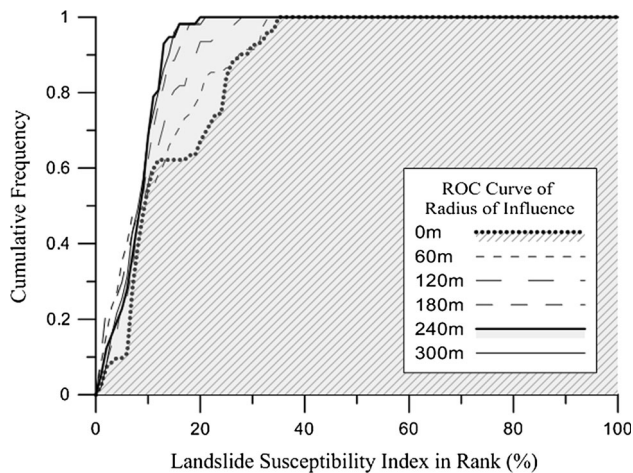
Fig. 6 AUCs of prediction maps according to the radius of influence (0, 60, 120, 180, 240, and 300 m)

value of AUC within a 0 m radius of influence (0.8649) remains valid, but LSI maps incorporating the optimized radius of influence will be more effective for hazard prediction.

Discussion

We compared both FR values that adopting the optimized radius of influence and not, in order to specify the reason of accuracy increase (Table 4). The number of classes which had comparatively constant FR values regardless of the radius of influence were 54 out of 85. From among these, six classes remained high in FR values, for instance >75.0 % in forest density and artificial hardwood forest in forest type. Still, eight classes kept 0 such as 40–50 years >50 % in forest age and wet land in land cover. The classes which retain their FR values, distributed dominantly around landslide occurred sites, can be divided into two categories, the spatial conditions where landslide can occur or not and triggering factors that can induce or prevent the landslide mechanism. It is the factors concerning with spatial conditions are >75.0 % in forest density and wet land in land cover. Meanwhile, artificial hardwood forest in forest type and 40–50 years >50 % in forest age contribute in triggering the landslides which directly affect to the ground stability and precipitation effects. These classes show the smaller fluctuation in FR values when varying radius of influence, which indicate the lower existence probability of other classes around adjacent areas, also assume rare influence to the adjacent areas.

The number of classes that the FR values were reversed below 1 to over 1, or vice versa, is 16 out of 85. However 6 out of 16 classes changed less than 200 %, which are titled as remained, and the other 10 classes showed sharp

variations. The 31 classes had their FR values increase or decrease over 200 %. It means that adjacent triggering factors concern with landslide occurrence as much as the gap grows bigger, far from 1 in FR value. 7 out of 85 classes whose FR values were zero, titled as advent, were additionally included in training domain when adopting the radius of influence. It is the benefit to apply the radius of influence when the spatial data are distributed excessively sparse or dense, especially for the classes whose FR values become closer to 1. The more optimizing the radius of influence, the more powerfully the error can be mitigated so that overestimation can be relieved. The correction effect is highly likely in case of higher spatial resolution. Unless, it is useless to widen the radius of influence because most properties of domain are in same classes. Therefore, it is the effective way to apply the radius of influence when whole spatial database has similar spatial resolution.

Conclusions

In this study, we applied the radius of influence to assessments of landslide susceptibility in order to statistically analyze the importance of various triggering factors. For the sake of compiling precise LSI maps, a series of spatial analysis models was built to find the optimum radius of influence for predicting landslide hazards with greater accuracy. We adopted the widely used FR method as the statistical method for each radius of influence, and selected 12 triggering factors from geographical and geological properties. The radius of influence was increased in 30-m increments from zero to 300 m, enabling us to construct 11 layered LSI maps. During the validation stage, we determined the optimum radius of influence of the target area to be 240 m by comparing the AUC values for each radius of influence.

The most striking effect of applying the appropriate radius of influence is the large reduction in data exaggeration and anomaly errors. As the domain of the training sites increases, the triggering factors in adjacent areas can be included in the statistical analysis. Subsequently, we discovered that several classes of triggering factors are either over- or underestimated, or even omitted, when the area of influence is not included. Conversely, if the radius of influence increases, the accuracy of the LSI map declines due to redundant referencing.

The most critical change of process was FR analysis when adopting the radius of influence. 36 % of whole classes varied in FR values over 200 % and even 7 classes whose value were zero gained positive values. As the amount of increase or decrease in FR values are greater, it is implied that the classes of adjacent area have more influence on landslide occurred sites. 64 % of classes were

Table 4 Changes in FR index (0 m and 240 m) for each class of triggering factor

Factor	Class	Radius of influence 0 m	Reversed 240 m	Remained	Advent	Increased over 200 %	Decreased over 200 %
Slope	0.00°–2.01°	0.041	0.230				√
	2.01°–5.05°	0.186	0.516				√
	5.05°–8.08°	0.779	1.039	√	√		
	8.08°–11.1°	1.499	1.499		√		
	11.1°–14.1°	2.054	1.757		√		
	14.1°–17.2°	2.773	2.048		√		
	17.2°–20.2°	3.390	1.826		√		
	20.2°–56.0°	2.414	1.548		√		
Aspect	Flat	0.082	0.356				√
	N	0.698	1.009	√	√		
	NE	0.909	0.864		√		
	E	1.289	0.827	√	√		
	SE	1.055	0.645	√	√		
	S	0.397	0.660		√		
	SW	0.272	0.785				√
	W	0.881	0.992		√		
	NW	1.327	1.500		√		
Curvature	Concave (−)	1.927	1.316		√		
	Flat	0.659	0.778		√		
	Convex (+)	1.271	1.395		√		
Land cover	Water	0.000	0.000		√		
	Urban area	0.148	0.535		√		
	Bare soil	0.198	0.498				√
	Wet land	0.000	0.000				√
	Grass land	0.582	0.516		√		
	Mountainous terrain	2.691	1.975		√		
	Rice field	0.818	0.851		√		
	Farm land	0.972	1.089		√		
Forest age (>50 % of whole timber)	0–10 years	19.64	2.629	√	√		
	10–20 years	1.839	2.384				√
	20–30 years	3.622	2.413		√		
	30–40 years	0.292	1.651		√		
	40–50 years	0.000	0.000	√			√
Forest density	<50.0 %	0.603	2.066		√		
	50.0–75.0 %	1.911	1.774	√			√
	>75.0 %	9.018	5.543		√		
Forest diameter	<6 cm	19.52	2.737		√		
	6–16 cm	3.162	2.453				√
	16–28 cm	0.143	0.828		√		
Forest type	Larix Kaempferi	0.000	0.000				√
	Mixed forest	0.306	1.240		√		
	Pinus Densiflora	2.990	2.484	√			√
	Hardwood forest	5.139	3.125		√		
	Farm land	0.113	0.333		√		
	Pinus Rigida	0.811	2.118				√
	Pinus Koraiensis	20.76	2.650	√			√

Table 4 continued

Factor	Class	Radius of influence 0 m	Reversed 240 m	Remained	Advent	Increased over 200 %	Decreased over 200 %
	Non-forest land	0.000	0.236				√
	Artificial hardwood forest	15.17	15.43			√	√
	Populus Deltoides	0.000	0.000		√		
Soil drainage (observation made by specialists' naked eye)	Excessively	2.209	2.056		√		
	Well	8.233	7.812		√		
	Moderately	0.437	0.519		√		
	Rarely	0.000	0.179		√		
	Poorly	0.186	0.378			√	√
	Outcrop	0.540	0.643				√
Soil material	Alluvium	0.024	0.132		√		
	Valley alluvium	0.000	0.092				√
	Gley alluvium	0.189	0.377			√	√
	Sedimentary soil	0.000	1.008		√		
	Red yellow soil	0.055	0.411	√		√	√
	Lithosol and red yellow soil	0.712	0.663				√
	Lithosol	2.490	2.325		√		
	Red yellow sedimentary soil	0.738	2.535		√		
	Outcrop	0.540	0.643	√			√
Soil texture	Coarse sand	0.012	0.096		√		
	Silt loam	2.122	2.771				√
	Gravelly loam	0.000	1.429		√		
	Clay sand	0.000	1.008	√		√	√
	Clay loam	0.229	0.228	√		√	√
	Clay	0.023	0.464		√		
	Gravelly sand	0.000	3.423				√
	Sandy silt loam	0.000	0.000	√		√	√
	Coarse loam	2.087	1.764		√		
	Rocky loam	0.738	2.535		√		
	Sandy loam	0.797	1.363	√			√
	Outcrop	0.540	0.643	√	√		
Soil thickness	>150 cm	0.059	0.325		√		
	125–150 cm	0.228	0.245				√
	100–125 cm	0.512	0.496		√		
	75–100 cm	2.122	2.771		√		
	50–75 cm	0.068	0.645		√		
	20–50 cm	1.897	1.793				√
	<20 cm	0.000	0.000		√		
	Outcrop	0.540	0.643		√		

independent from the radius of influence, especially whose spatial data were distributed relatively flat. It is less likely for the classes to concern with triggering landslide from the adjacent areas.

The classes with extreme FR values, such as wet land in land cover and >75.0 % in forest density, are highly related with geological properties where the landslide events occur, while 40–50 years >50 % in forest age and artificial

hardwood forest in forest type are triggering conditions of landslide, especially with soil failure. 7 Classes which had rapid changes in FR values had been considered as extreme triggering factors without the radius of influence but not actually. It is the reason that the landslides spread and damage adjacent areas why the radius of influence should be considered to include those factors in the stage of statistical analysis. As a result, an LSI map that considers the optimum radius of influence is 5.95 % points more accurate than the map without. We anticipate finding more suitable values for the radius of influence in future studies.

The peak AUC value in our experiment is clearly apparent, and we stress that it is important to find the point at which AUC values cease to rise. If the radius of influence is set excessively high or low, the accuracy of the LSI maps will be reduced. The optimum radius of influence differs case by case. The number of sections and range of the radius of influence must be established in a manner that reflects each specific situation. If the spatial resolution of the database is insufficient or excessively fine, the area of influence might not be detected. With this in mind, it is prudent to apply the radius of influence to statistical models that incorporate not only FR analysis but also other methods, such as analytical hierarchy processes, bivariate statistics, logistic regression, and artificial neural networks.

References

Akgun A, Türk N (2010) Landslide susceptibility mapping for Ayvalik (Western Turkey) and its vicinity by multicriteria decision analysis. Environ Earth Sci 61(3):595–611

Choi J, Oh HJ, Lee HJ, Lee C, Lee S (2012) Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. Eng Geol 124:12–23

Goltz C (1996) Multifractal and entropic properties of landslides in Japan. Geol Rundsch 85:71–84

Guzzetti F, Peruccacci S, Rossi M, Stark CP (2007) Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorol Atmos Phys 98:239–267

Hong SY, Zhang YS, Hyun BK, Sonn YK, Kim YH, Jung SJ, Park CW, Song KC, Jang BC, Choe EY, Lee YJ, Ha SK, Kim MS, Lee JS, Jung GB, Ko BG, Kim GY (2009) An introduction of Korean soil information system. Korean J Soil Sci Fert 45(1):21–28

Jones KE, Preston NJ (2012) Spatial and temporal patterns of off-slope sediment delivery for small catchments subject to shallow landslides within the Waipaoa catchment, New Zealand. Geomorphology 141–142:150–159

Kim KH, Jung HR, Park JH, Ma HS (2011) Analysis on rainfall and geographical characteristics of landslides in Gyeongnam Province. J Korean Environ Restor Technol 14:33–45

Kim HG, Lee DK, Park C, Kil S, Son Y, Park JH (2015) Evaluating landslide hazards using RCP 4.5 and 8.5 scenarios. Environ Earth Sci 73:1385–1400

Lee S, Lee M (2005) Detecting landslide location using KOMPSAT 1 and its application to landslide-susceptibility mapping at the Gangneung area, Korea. Adv Space Res 38:2261–2271

Lee S, Min K (2001) Statistical analysis of landslide susceptibility at Yongin, Korea. Envrion Geol 40:1095–1113

Lee S, Ryu JH, Lee MJ, Won JS (2006) The application of artificial neural networks to landslide susceptibility mapping at Janghung, Korea. Math Geol 38(2):199–220

Lee S, Hwang J, Park I (2013) Application of data-driven evidential belief functions to landslide susceptibility mapping in Jinbu, Korea. Catena 100:15–30

Liucci L, Melelli L, Suteanu C (2015) Scale-invariance in the spatial development of landslides in the Umbria Region (Italy). Pure Appl Geophys 172(7):1959–1973

Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004) Landslide inventories and their statistical properties. Earth Surf Process Landf 29:687–711

Mason SJ, Graham NE (2002) Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: statistical significance and interpretation. Q J R Meteorol Soc 128:2145–2166

Massey CI, Petley DN, McSaveney MJ (2013) Patterns of movement in reactivated landslides. Eng Geol 159:1–19

Oh HJ, Pradhan B (2011) Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. Comput Geosci 37(9):1264–1276

Park S, Choi C, Kim B, Kim J (2013) Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. Environ Earth Sci 68(5):1443–1464

Petley DN, Mantovani F, Bulmer MH, Zannoni A (2005) The use of surface monitoring data for the interpretation of landslide movement patterns. Geomorphology 66:133–147

Pouduyal CP, Chang C, Oh HJ, Lee S (2010) Landslide susceptibility maps comparing frequency ratio and artificial neural networks: a case study from the Nepal Himalaya. Environ Earth Sci 61(5):1049–1064

Pradhan B (2011) Use of GIS-based fuzzy logic relations and its cross application to produce landslide susceptibility maps in three test areas in Malaysia. Environ Earth Sci 63(2):329–349

Reis S, Yalcin A, Atasoy M, Nisanci R, Bayrak T, Erduran M, Sancar C, Ekercin S (2012) Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio and analytical hierarchy methods in Rize province (NE Turkey). Environ Earth Sci 66:2063–2073

Schleier M, Bi R, Rohn J, Ehret D, Xiang W (2014) Robust landslide susceptibility analysis by combination of frequency ratio, heuristic GIS-methods and ground truth evaluation for a mountainous study area with poor data availability in the Three Gorges Reservoir area, PR China, 71:3007–3023

Shahabi H, Hashim M, Ahmad BB (2015) Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central Zab basin, Iran. Environ Earth Sci 73:8647–8668

Suh J, Choi Y, Roh TD, Lee HJ, Park HD (2011) National-scale assessment of landslide susceptibility to rank the vulnerability to failure of rock-cut slopes along expressways in Korea. Environ Earth Sci 63(3):619–632

Suh J, Choi Y, Park HD, Yoon SH, Go WR (2013) Subsidence hazard assessment at the Samcheok Coalfield, South Korea: a case study using GIS. Environ Eng Geosci 19(1):69–83

Ulrich K, Growley BJ, Khattak GA, Owen LA (2008) GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. Geomorphology 101(40):631–642

Van Westen CJ, Rengers N, Soeters S (2003) Use of geomorphological information in indirect landslide susceptibility assessment. Nat Haz 30(3):399–419

Yalcin A, Reis S, Aydinoglu AC, Yomralioglu T (2011) A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. Catena 85:274–287

Yilmaz I (2009) Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides. Comput Geosci 35(6):1125–1138

Yilmaz I (2010) Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. Environ Earth Sci 61(4):821–836

Yilmaz I, Topal T, Süzen ML (2012) GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). Environ Earth Sci 65(7):2161–2178

Yune CY, Chae YK, Paik J, Kim G, Lee SW, Seo HS (2013) Debris flow in metropolitan area—2011 Seoul debris flow. J Mt Sci 10(2):199–206