

TPI와 경사도 조합을 이용한 산사태 위험도 평가 Landslide Susceptibility Assessment Using TPI-Slope Combination

이한나¹⁾ · 김기홍²⁾
Lee, Han Na · Kim, Gihong

Abstract

TSI (TPI-Slope Index) which is the combination of TPI (Topographic Position Index) and slope was newly proposed for landslide and applied to a landslide susceptibility model. To do this, we first compared the TPIs with various scale factors and found that TPI350 was the best fit for the study area. TPI350 was combined with slope to create TSI. TSI was evaluated using logistic regression. The evaluation showed that TSI can be used as a landslide factor. Then a logistic regression model was developed to assess the landslide susceptibility by adding other topographic factors, geological factors, and forestal factors. For this, landslide-related factors that can be extracted from DEM (Digital Elevation Model), soil map, and forest type map were collected. We checked these factors and excluded those that were highly correlated with other factors or not significant. After these processes, 8 factors of TSI, elevation, slope length, slope aspect, effective soil depth, tree age, tree density, and tree type were selected to be entered into the regression analysis as independent variables. Three models through three variable selection methods of forward selection, backward elimination, and enter method were built and evaluated. Selected variables in the three models were slightly different, but in common, effective soil depth, tree density, and TSI was most significant.

Keywords : Topographic Position Index, TPI-Slope Index, Landslide Susceptibility, Logistic Regression Model

초 록

TPI (Topographic Position Index)와 경사도를 조합하여 새로운 산사태 인자인 TSI (TPI-Slope Index)를 제안하고 산사태 예측모형에 적용하였다. 이를 위해, 먼저 다양한 분석 반경의 TPI를 서로 비교하여 TPI350이 연구 대상 지역에 가장 적합함을 알아내었고, 이를 경사도와 조합하여 TSI를 제작하였다. 본 논문에서 제안한 TSI의 적용성을 평가하기 위해 로지스틱 회귀분석을 이용한 결과, 산사태 예측 모형에 활용할 수 있다는 결론을 얻었다. 그 후, 기타 지형 정보들과 토양 및 임상 정보를 추가하여 산사태 위험도를 평가하는 로지스틱 회귀 모형을 제작하였다. 이를 위해 DEM (Digital Elevation Model), 토양도, 임상도로부터 추출할 수 있는 산사태 관련 인자들을 수집하고 이들을 검토하여 다른 인자와 상관도가 높거나 산사태와의 연관성이 낮은 인자들은 우선 제외하였다. 그 결과, TSI, 고도, 사면 길이, 경사향, 유효 토심, 영급, 나무 밀도, 임상 등 8개의 인자가 선정되어 회귀분석에 독립변수로 입력되었다. 변수의 입력 방법(전진 선택법, 후진 제거법, 직접 선택법)에 따라 3가지 모형을 생성하였고, 이들에 대한 평가를 수행하였다. 세 모형에서 선택된 변수는 조금씩 다르지만, 공통적으로 유효 토심, 나무 밀도, TSI 인자의 중요도가 높은 것으로 나타났다.

핵심어 : 지형위치지수, 지형위치-경사 지수, 산사태 위험도, 로지스틱 회귀 모형

Received 2018. 11. 11, Revised 2018. 12. 04, Accepted 2018. 12. 18

1) Member, Industry-Academia Collaboration Foundation, Gangneung-Wonju National University (E-mail: leehn77@hanmail.net)

2) Corresponding Author, Member, Dept. of Civil Engineering, Gangneung-Wonju National University (E-mail: ghkim@gwnu.ac.kr)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

1.1 연구 배경

TSI (TPI-Slope Index)는 Guisan *et al.* (1999)이 식물종 분포에 대한 연구를 하면서 처음 제안한 지형 특성으로, DEM (Digital Elevation Model)에서 어떤 픽셀의 고도값과 주변 픽셀들의 평균 고도값의 차이를 나타내는 값이다. 여기서 주변의 범위 즉, 분석 반경을 어떻게 정의하는가에 따라 TPI 값이 달라지며, 이에 따라 지형 분류도 달라질 수 있다. 때문에 일반적으로 TPI300, TPI2000처럼 뒤에 분석 반경을 나타내는 숫자를 붙여서 표기한다. TPI300은 분석 반경이 300m란 뜻이다. TPI는 계산이 쉽고 지형의 특징을 직관적으로 표현해준다는 큰 장점이 있어서 산사태 인자로 매력적이다. TPI에 대한 자세한 내용은 Weiss (2001)의 연구에서 소개되었다. Weiss는 TPI300과 TPI2000을 결합하여 지형을 10가지로 분류하였는데, 국내에서도 이러한 지형분류법을 토대로 연구가 진행되었다. Woo *et al.* (2008)은 국내 지형 특색을 고려하여 TPI100과 TPI500을 결합하여 지형을 분류하였다. Jang *et al.* (2009)은 3곳의 연구대상 지역을 선정하고 각 지역의 지형 기복량의 변화를 분석하여 각각 TPI50과 TPI250, TPI50과 TPI350, TPI50과 TPI400을 결합하여 지형을 분류하였다. Lee *et al.* (2015)도 TPI를 포함한 세 가지 지형분류기법을 비교하면서, TPI60과 TPI300, TPI60과 TPI340, TPI80과 TPI460 등의 조합으로 지형을 분류하였다. 이와 같이 두 가지 분석 반경을 갖는 TPI를 결합하여 지형을 분류하는 이유는 큰 계곡 안의 작은 언덕이나 고산지대의 배수로 등과 같은 다양한 지형을 구분하기 위해서이다. Weiss (2001)는 TPI의 활용 분야를 토사 이동, 국지적 물수지 분석에서부터 기후, 바람길, 생태 특성에 대한 연구까지 매우 넓게 보았다. 때문에 거시적 지형과 세부 지형을 모두 고려한 지형 분류기법을 제시한 것이다. 그러나 연구 주제가 산사태로 한정된다면 Weiss의 분류기법을 그대로 적용하기에는 무리가 있는데, 우리나라의 산지에서 큰 계곡 안의 작은 언덕이나 고산지대의 배수로와 같은 지형은 거의 없기 때문이다. Woo *et al.* (2008)와 Jang *et al.* (2009)의 연구를 살펴보면 국내 대상지역에서 큰 계곡 안의 작은 언덕이나 고산지대의 배수로로 분류되는 지역의 비중이 0.0%에서 0.2% 사이로 사실상 분류할 필요가 없다는 것을 알 수 있다. 본 연구에서는 어떤 지점의 거시적이면서 세부적인 지형 특성을 알아내는 데에 중점을 둔 기존 연구와는 달리, 경사면에서 해당 지점의 위치를 특정 하는데 초점을 맞추어 TPI의 새로운 활용 방법을 제안하고자 한다.

새롭게 제안된 인자를 포함한 산사태 관련 인자들의 평가에는 로지스틱 회귀모형을 활용하였다. 로지스틱 회귀모형은 의

학, 금융, 범죄학, 생태학, 언어학, 공학 등 다양한 분야에 활용되면서 발전해왔다(Hosmer and Lemeshow, 2000). 이 과정에서 산사태 관련 연구에도 많이 응용되었고 좋은 성과를 내어왔다. 로지스틱 회귀모형의 큰 장점 중 하나는 다양한 형태의 독립 변수를 쉽게 처리할 수 있다는 점일 것이다(Costanzo *et al.*, 2014). 특히 관련 소프트웨어의 발달로, 방대한 자료를 짧은 시간에 간편하게 처리하고 결과를 확인할 수 있게 되어 활용도가 더욱 높아지고 있다. 국내에서도 로지스틱 회귀모형을 이용한 산사태 연구가 다수 수행되었다. Chae *et al.* (2004)은 5개의 지질 및 토질 인자와 2개의 지형인자를 이용하여 로지스틱 회귀모형을 개발하였으며, Lee *et al.* (2006)은 총 6개의 지형, 토양 및 환경 인자에 로지스틱 회귀분석과 AHP 기법을 적용하였다. Lee and Kim (2012)은 지형, 임상, 토양 자료를 이용하여 로지스틱 회귀모형을 개발하였다.

1.2 연구 대상 지역 및 연구 방법

본 연구는 강원도 평창군 진부면 봉산리 봉산천 주변의 산지에서 발생한 산사태와 주변 지형을 대상으로 수행되었다. 이 지역의 정사영상과 DEM을 분석하여 Fig. 1과 같이 97개의 산사태 발생 지점을 추출하였으며, 산사태가 발생하지 않은 396 지점을 추가로 생성하였다. 사진 검토 작업에는 493점 모두를 활용하였고, 회귀분석에는 발생 지점과 비발생 지점의 비율을 맞추기 위해(Costanzo *et al.*, 2014) 비발생 지점 중 296 지점을 제외한 197개 지점만을 사용하였다.

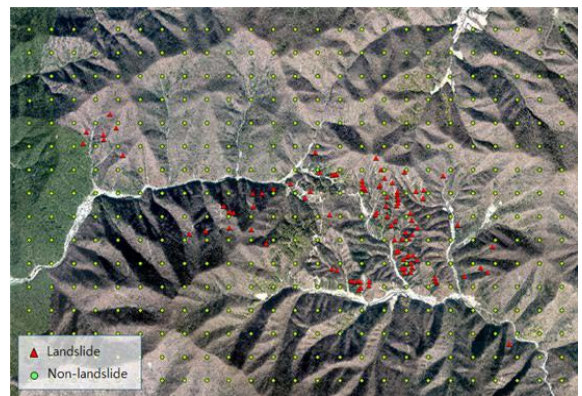
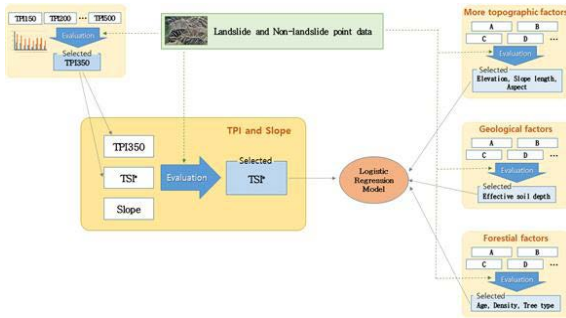


Fig. 1. Study area and landslide/non-landslide points

Fig. 2에 연구 흐름을 정리하였다. TPI와 경사도를 조합하여 새로운 인자인 TSI를 제안하였고, 여기에 기타 지형인자와 토양 및 임상 인자를 추가하여 산사태 예측모형을 제작하였다. 산사태 관련 인자들은 이항 로지스틱 회귀분석에 평가 대상 인자들

독립변수로, 산사태 발생 여부를 종속변수로 입력하여 제작된 산사태 예측모형의 적합성을 참고하여 평가되었다.

다수 모형의 적합성을 비교할 때, 자유도가 같은 경우에는 주로 $-2\text{Ln}\mathcal{L}$ 값을 참고하였다. $-2\text{Ln}\mathcal{L}$ 는 로그 우도함수에 -2 를 곱한 값으로, 0에서 무한대까지의 값을 가지며 모형의 적합도가 좋을수록 0에 가까운 값을 갖게 된다(Hong, 2005). 모형의 자유도가 서로 다른 경우에는 Akaike 정보지수(AIC: Akaike Information Criterion)와 Bayesian 정보지수(BIC: Bayesian Information Criterion)를 사용하였다(Akaike, 1973; Raftery, 1995). 이들은 회귀 모형의 비교에 널리 사용되는 지수로, 모두 로그 우도함수를 이용하여 모형의 적합도를 반영하면서 모형의 미지수 개수 또는 자유도를 고려하여 간명성도 함께 평가한다. AIC, BIC 모두 값이 작을수록 좋은 모형이다(Hong, 2005).



*TSI : TPI-Slope Index

Fig. 2. Flow chart

경우에 따라 분류 정확도도 검토하였다. 본 논문에서 사용한 분류 정확도는 정밀도(고위험으로 예측된 지점에서 실제 산사태가 발생한 확률, precision), 검출률(산사태가 실제 발생한 지점을 고위험 지역으로 분류한 확률, recall), 전체 정확도(전체 지점중 제대로 분류된 지점의 비율, overall), 세 가지이다. 산사태는 심각한 피해를 가져올 수 있는 자연 재해이므로 AIC와 BIC 평가 결과가 서로 엇갈리는 경우와 같이 추가적인 판단 기준이 필요할 경우, 검출률에 좀 더 무게를 두었다.

2. TPI와 경사도

2.1 TPI 선정

TPI 값의 범위가 분석 반경에 따라 달라지는 문제를 해결하기 위해 값을 표준화시켜 사용하는 경우가 많다. 본 연구에서도 표준화된 TPI를 사용하였으며, 이후 언급되는 TPI는 모두

표준화된 것이다. 먼저, 어떤 분석 반경의 TPI가 연구 대상 지역에 적합한지 알아내기 위해 분석 반경 150m에서부터 500m까지, 50m 간격의 TPI를 제작하였다. TPI는 연속형 자료이지만 Table 1과 같이 5개의 등급으로 나누어 범주형 변수로 변환하였다. TPI 값이 -0.5σ 에서 0.5σ (σ =표준편차) 사이이면 평면 또는 사면의 중심부로 볼 수 있다(Weiss, 2001).

Table 1. TPI classification

| Standardized TPI | $-\infty \sim -1.5$ | $-1.5 \sim -0.5$ | $-0.5 \sim 0.5$ | $0.5 \sim 1.5$ | $1.5 \sim \infty$ |
|------------------|---------------------|------------------|-----------------|----------------|-------------------|
| TPI class | T1 | T2 | T3 | T4 | T5 |

다양한 분석 반경의 8가지 TPI는 이항 로지스틱 회귀분석에 각 TPI 등급을 독립변수로, 산사태 발생 여부를 종속변수로 입력하여 제작된 산사태 예측모형의 $-2\text{Ln}\mathcal{L}$ 값을 기준으로 평가하였다. 각 TPI 등급만을 독립변수로 사용한 모형과 경사도를 추가하여 두 개의 독립변수를 사용한 모형의 평가 결과를 Fig. 3에 정리하였다. 두 경우 모두에서 분석 반경이 150m보다 넓어질수록 $-2\text{Ln}\mathcal{L}$ 값이 작아지다가 350m에서 최소가 되고 이후 다시 커지는 추세를 볼 수 있다. 따라서 본 연구에서는 TPI350을 사용하였다.

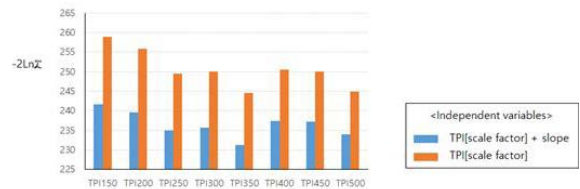


Fig. 3. $-2\text{Ln}\mathcal{L}$ of each logistic regression model

2.2 TPI-경사도 조합과 TSI

Fig. 4는 연구 대상 지역의 한 경사면에서 고도와 TPI350, 경사도의 변화를 보여준다. TPI는 평지와 경사면의 중심부에서 동일하게 0에 가까운 값을 갖기 때문에 TPI만으로는 해당 지점의 지형을 구분할 수 없다. 또, 경사도만으로는 해당 지점이 완만한 경사의 중심부인지, 급경사면의 상단 혹은 하단인지 구별할 수 없다. 하지만, Fig. 4를 살펴보면 TPI와 경사도의 조합을 이용하여 지형을 보다 정확히 구분할 수 있음을 알 수 있다. 예를 들어, 경사면의 중심부에서는 경사도가 비교적 높은 값을 갖는 반면 TPI는 0에 가까운 값을 갖는다. 경사면의 하단부에서는 경사도가 낮아지면서 TPI는 음의 값을 갖는다. 따라서 우리는 TPI-경사도 조합을 이용한 새로운 지표를 만들고, 이를 산사태

위험도 예측 모형의 지형 인자로 활용하기로 하였다.

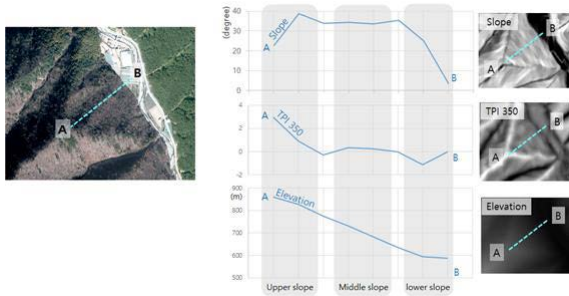


Fig. 4. Variation of slope, TPI350, and elevation at a slope

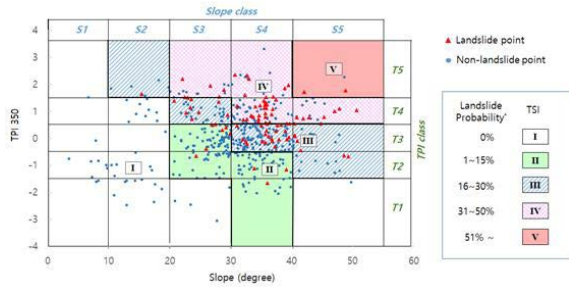


Fig. 5. Distribution of landslide/non-landslide points in slope-TPI350 plane and TSI

Fig. 5는 가로축을 경사도, 세로축을 TPI350으로 하는 좌표 평면에 97개 산사태 발생 지점과 396개 미발생 지점의 분포를 나타낸 그래프이다. TPI350이 -0.5~1.5이고, 경사도가 30~40 도인 구간에 산사태 발생지점이 밀집되어 있음을 볼 수 있다. 앞서 설명한 바와 같이 TPI350은 T1~T5의 5개 등급으로 나누었다(Table 1). 마찬가지로 경사도 또한 10°간격으로 S1~S5의 5개 등급으로 나누었다. 이들을 조합하면 25개의 TPI-경사도 조합이 만들어진다. TPI350이 0에 가깝고 경사도가 낮은 T3S1에

해당되는 지점은 평면, TPI350이 0에 가깝지만 경사도가 높은 T3S4, T3S5는 경사면의 중간 지점이다. T5S3은 사면의 윗부분, T1S3은 사면의 아랫부분으로 볼 수 있다.

25개 TPI-경사도 각 영역에서 Eq.(1)과 같이 산사태 발생 확률을 계산하였다.

$$\text{Landslide probability} = \frac{\text{Number of landslide points}}{\text{Total number of points}} \quad (1)$$

이 확률을 기준으로 Fig. 5와 같이 TPI-Slope Index(TSI)를 결정하였다. 예를 들어, T3S4에 해당하는 지점 수는 총 114개 인데 이 중 산사태가 발생한 지점은 19 지점으로 16.7%의 발생 확률을 보이며, 따라서 TSI는 'III'으로 분류된다. T3S4와 경사도는 같지만 TPI350 값은 더 큰 T4S4의 경우 총 68 지점 중 산사태 발생 지점이 27 지점으로, 발생 확률이 39.7%이므로, TSI는 'IV'로 분류된다.

2.3 TSI의 적합성

각 지점에 부여된 TSI를 산사태 예측을 위한 회귀분석에 사용하는 것이 적합한가를 판단하기 위해 TPI350과 경사도, 그리고 TSI를 독립변수로 사용한 각 회귀 모형(각각 Table 2의 모형 A, B, E)을 비교하였다. 이와 더불어 보다 정확한 평가를 위해 TPI350과 경사도를 함께 사용한 모형(모형 C), TSI를 구분하지 않고 T3S1, T5S3와 같은 TPI-경사도 조합을 바로 독립 변수로 사용한 모형(모형 D)도 함께 비교하였다. 비교의 척도로 모형의 $-2\ln L$ 값과 AIC, BIC를 사용하였다. AIC, BIC는 모형 D와 E만 계산하였다. 자유도가 가장 낮은 모형 E보다 $-2\ln L$ 가 큰 모형은 더 이상 낫할 필요가 없기 때문이다. 독립변수가 하나뿐인데도 자유도가 4, 8과 같은 값을 갖는 것은 범주형 변수의 경우 (범주의 수-1) 만큼의 더미변수가 생기기 때문이다. TPI350은 5개의 범주로 구성된 범주형 변수이므로 자유도가 4이다.

Table 2. Evaluation of the factors: TPI350, slope, and TSI

| Model | Independent variables in the model | Degree of freedom | $-2\ln L$ | AIC* | BIC** | Percentage correct(%) | | |
|-------|------------------------------------|-------------------|-----------|--------|--------|-----------------------|--------|---------|
| | | | | | | Precision | Recall | Overall |
| A | TPI350 | 4 | 244.645 | - | - | 66.3 | 64.9 | 66.5 |
| B | Slope | 4 | 257.378 | - | - | 55.1 | 77.3 | 57.9 |
| C | TPI350 and slope | 8 | 231.245 | - | - | 68.4 | 69.1 | 69.0 |
| D | TPI-slope combination | 13 | 226.063 | 254.06 | 21.69 | 68.3 | 71.1 | 69.5 |
| E | TPI-slope index (TSI) | 4 | 231.036 | 241.04 | -20.89 | 61.8 | 91.8 | 68.0 |

* AIC = $-2\ln L + 2q$, where q: the number of parameters in the model (Akaike, 1973)

** BIC = $-(2\ln L_0 - (2\ln L_M)) + df \cdot \ln N$, where L_0 : likelihood of the null model, L_M : likelihood of the model, df = degree of freedom, N: sample size (Raftery, 1995)

Table 3. Evaluation of topographic factors

| | Degree of Freedom | χ^2 | P-value | $-2\ln \mathcal{L}$ | AIC | BIC |
|----------------------|-------------------|----------|---------|---------------------|---------|-----------|
| Elevation | 1 | 22.038 | 0.000 | 251.016 | 255.016 | -16.7548 |
| Slope length | 1 | 23.652 | 0.000 | 249.402 | 253.402 | -18.3688 |
| Curvature plan | 1 | 1.061 | 0.303 | 271.993 | - | - |
| Curvature profile | 1 | 2.827 | 0.093 | 270.227 | - | - |
| Curvature general | 1 | 2.577 | 0.108 | 270.477 | - | - |
| Aspect(8 directions) | 7 | 32.646 | 0.000 | 240.408 | 256.408 | 4.3364261 |

□ : Excluded from regression

$-2\ln \mathcal{L}$ 가 가장 작은 모형은 D 모형이지만, AIC, BIC를 통해 모형 E의 적합도가 가장 좋은 것을 알 수 있다. BIC의 차이가 10 이상이면 모형의 적합도에 매우 큰 차이가 있다고 볼 수 있다(Raftery, 1995). 분류 정확도를 살펴보면, 전체 정확도는 모형 E가 모형 C, D에 비해 다소 낮지만, 검출률은 모형 E가 가장 뛰어나다.

3. 기타 산사태 관련 인자

3.1 기타 지형 인자

TPI와 경사도를 제외한 기타 지형인자로 고도, 사면 길이, 곡률(평면곡률, 단면곡률, 총곡률), 경사향(8방향)을 고려하였다. 이들에 대한 평가는 Table 3과 같다. 세 가지 곡률은 χ^2 검증에서 유의확률이 5%를 넘는데, 이는 통계적으로 의미가 없다는 뜻이므로 모형에서 제외하였다. 타 연구에서도 곡률이 모형의 성능을 저하시킨다는 이유로 제외된 경우가 있다(Othman *et al.*, 2015).

3.2 토양 인자

토양도에서 얻을 수 있는 토양인자는 심토 토성, 토양 구조, 심토 자갈함량, 모재, 배수 등급, 유효 토심, 침식 등급, 표토 토성, 토양목, 토양아목, 토양통, 토양 유형 등 매우 다양하다. 이 중 토양목, 토양아목, 침식 등급은 연구대상지역 전체에서 한 가지 값을 가지므로 연구에서 제외하였다. 심토 토성과 토양 유형 역시 계곡과 그 외 지역을 구분하는 수준의 분포를 보여(Fig. 6) 분석 가치가 없다고 판단하였으므로 연구에서 제외하였다. Fig. 7을 보면, 토양 구조는 범주가 다양하지만 대상 지역 내에서는 세 가지 값만 가지며, 심토 자갈함량과 동일한 분포를 보인다. '999'는 픽셀값이 없음을 의미한다. 동일한 분포를 보이는 두 인자 중 범주의 수가 더 많은 토양구조를 연구에서 제외하였다.

심토 자갈함량, 토양통, 유효 토심, 토양 모재, 표토 토성, 배수 등급을 평가한 결과는 Table 4와 같다. 서로 큰 차이가 없

는 가운데 유효 토심의 AIC, BIC가 가장 작다. Fig. 8을 보면 토양통과 표토 토성은 거의 동일한 분포를 보이며, 유효 토심, 토양 모재, 배수 등급은 토양통이 단순화되는 형태로 나타난다. 이들 인자의 상관 분석 결과 상관계수도 높게 나타났다. 따라서 유효 토심을 제외한 토양 인자는 분석에서 제외하였다.

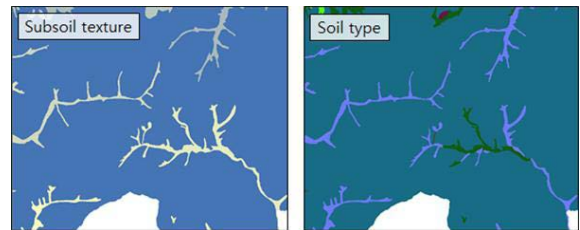


Fig. 6. Geological factors - Subsoil texture and soil type

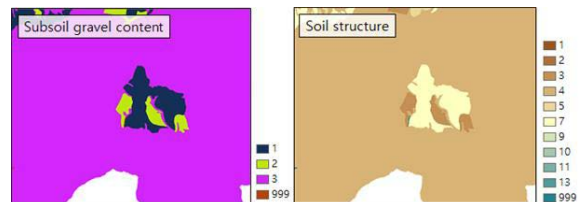


Fig. 7. Geological factors - Subsoil gravel content and soil structure

Table 4. Evaluation of geological factors

| | Degree of Freedom | $-2\ln \mathcal{L}$ | AIC | BIC |
|------------------------|-------------------|---------------------|---------|----------|
| Soil series | 3 | 189.736 | 197.736 | -67.4694 |
| Effective soil depth | 1 | 193.079 | 197.079 | -74.6918 |
| Parent material | 1 | 196.894 | 200.894 | -70.8778 |
| Topsoil texture | 2 | 192.875 | 198.875 | -69.6126 |
| Drainage | 1 | 203.992 | 207.992 | -63.7788 |
| Subsoil gravel content | 2 | 191.538 | 197.538 | -70.9506 |

□ : Excluded from regression

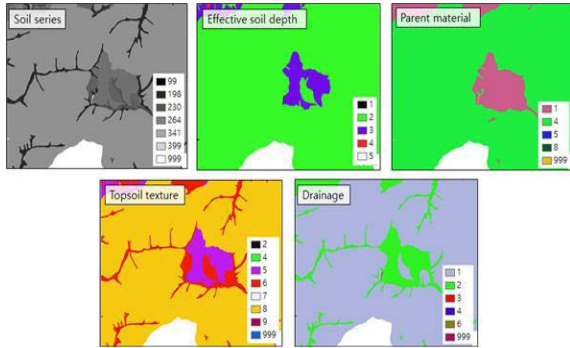


Fig. 8. Geological factors – Soil series, effective soil depth, parent material, topsoil texture, and drainage

3.3 임상 인자

임상 인자로는 영급(age), 경급(diameter), 밀도, 나무 종류가 있다. Fig. 9에서 보듯이 영급과 경급은 분포가 거의 동일하며, 분석 결과 이 둘 사이의 피어슨 상관 계수는 0.97로 나타났다.

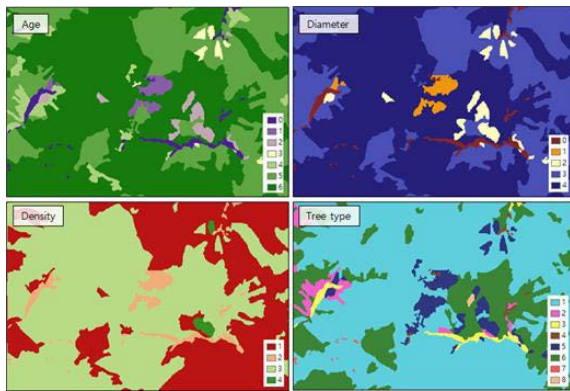


Fig. 9. Forestal factors

Table 5. Evaluation of forestial factors

| | Degree of Freedom | $-2\ln\mathcal{L}$ | AIC | BIC |
|-----------|-------------------|--------------------|---------|----------|
| Age | 6 | 223.611 | 237.611 | -17.7438 |
| Diameter | 4 | 232.419 | 242.419 | -19.5032 |
| Density | 3 | 244.419 | 252.419 | -12.7854 |
| Tree type | 5 | 234.196 | 246.196 | -12.442 |

□ : Excluded from regression

각 임상 인자에 대한 평가는 Table 5와 같다. 영급과 경급을 비교하면, AIC를 기준으로 할 경우 영급이 낮고 BIC를 기준으로 평가하면 경급이 낮다. 본 연구에서는 모형의 간명성보다 $-2\ln\mathcal{L}$ 값으로 설명되는 모형의 예측력을 더 중요한 지표로 간주하여 영급을 선택하였으며 영급과 상관관계가 매우 높은 경급은 고려 대상에서 제외하였다.

4. 산사태 예측 모형

사전 검토 과정에서 적합하지 못한 것으로 분류된 인자들을 제외한 8개 인자(TSI, 고도, 사면길이, 경사향, 유효 토심, 영급, 나무 밀도, 임상)를 독립변수로 입력하여 이항 로지스틱 회귀분석을 수행하였다. 변수 입력 방법에 따라 3가지 회귀모형이 제작되었다.

첫 번째는 우도비 검증을 통해 가장 설명력이 높은 변수를 차례로 모형에 포함시키는 전진 선택방법에 의한 모형(M1), 두 번째는 먼저 모든 변수를 포함시키고 중요하지 않은 변수를 제거해 나가는 후진 제거방법에 의한 모형(M2)이다. 이 때, 변수를 하나씩 추가 또는 제거할 때마다 모형의 개선 정도를 평가하고 더 이상 유의미한 개선이 이루어지지 않으면 그 단계에서 작업을 종료하게 된다. 마지막 세 번째는 직접 입력 방법을 사용하

Table 6. Performance of 3 logistic regression models

| Model | Method | $-2\ln\mathcal{L}$ | χ^2 | Percentage correct(%) | | | AIC | BIC |
|-------|---|--------------------|----------|-----------------------|--------|---------|-------|-------|
| | | | | Precision | Recall | Overall | | |
| DoF* | Variables | | | | | | | |
| M1 | Forward | | | | | | | |
| 7 | Effective soil depth, tree density, TSI | 146.0 | 127.1 | 91.2 | 75.3 | 84.3 | 162.0 | -90.1 |
| M2 | Backward | | | | | | | |
| 8 | Effective soil depth, tree density, TSI, elevation | 142.2 | 130.8 | 89.9 | 72.3 | 82.7 | 160.2 | -88.6 |
| M3 | Enter | | | | | | | |
| 25 | Effective soil depth, tree density, TSI, elevation, slope length, aspect, tree age, tree type | 118.1 | 154.9 | 88.0 | 83.5 | 86.3 | 170.1 | -22.8 |

* DoF : Degree of Freedom

여 8개의 인자를 모두 변수로 사용한 모형(M3)이다.

전진 선택법을 사용한 경우(M1)에는 유효 토심, 나무 밀도, TSI 등 3개의 변수가 모형에 포함되었다. 후진 제거법을 사용한 경우(M2)에는 8개의 변수 중 사면길이, 영급, 경사향, 임상 순으로 제거되어 유효 토심, 나무 밀도, TSI, 고도가 모형에 포함되었다. 세 모형에 대한 간단한 설명과 평가를 Table 6에 정리하였다.

모형의 예측력만 고려한다면 모든 변수가 포함된 모형 M3의 성능이 가장 뛰어나다. 그러나 모형의 간명성을 함께 고려한다면 AIC, BIC 값이 더 작은 모형 M1, 모형 M2가 더 뛰어나다고 할 수 있다.

5. 결론

본 연구에서는 산사태 인자로서 TPI와 경사도가 갖고 있는 단점을 보완하기 위해 이 둘의 조합으로 만들어지는 TSI를 제안하였다. 로지스틱 회귀분석을 이용하여 TPI, 경사도, TSI를 평가한 결과, TSI를 사용한 모형 E의 성능이 TPI 또는 경사도를 단독으로 사용한 모형 A, B 보다 월등히 좋았다. TPI와 경사도, 두 독립 변수를 사용한 모형 C과 비교해보면, 모형 E가 모형 C보다 $-2\ln\mathcal{L}$ 값이 약간 낮았다. 모형 C의 자유도가 모형 E보다 매우 높기 때문에 모형 E는 모형 C보다 성능이 좋다고 판단된다. TSI를 결정하기 전 단계인 TPI-경사도 조합을 독립변수로 사용한 모형 D는 모형 E보다 $-2\ln\mathcal{L}$ 값이 낮았다. 그러나 자유도가 매우 높아서 AIC, BIC 값은 모형 E보다 현저히 크다. 따라서 TSI를 적용한 모형의 성능이 가장 높았다고 판단할 수 있으며 향후 산사태 예측을 위한 입력인자로 TSI를 적용하는 것이 위험도를 산정의 정확도를 높이는 데 도움이 될 수 있을 것으로 판단된다. 본 논문에서 수행한 로지스틱 회귀 분석에 따르면, 산사태 발생과 연관이 있다고 추정되는 다양한 인자들 중에서 유효 토심, 나무 밀도, TSI가 산사태 위험도 평가에서 가장 중요한 인자이고, 다음으로 고도가 영향력이 큰 것으로 나타났다.

감사의 글

이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(NRF-2017R1D1A1B06036077).

References

Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, *Proceedings of the*

second International Symposium on Information Theory, Akademiai Kiado, 2-8 September, Tsahkadsor, Armenia, pp. 267-281.

Chae, B.G., Kim, W.Y., Cho, Y.C., Kim, K.S., Lee, C.O., and Choi, Y.S. (2004), Development of a logistic regression model for probabilistic prediction of debris Flow, *The Journal of Engineering Geology*, Vol. 14, NO. 2, pp. 211-222. (in Korean with English abstract)

Costanzo, D., Chacón, J., Conoscenti, C., Irigaray, C., and Rotigliano, E. (2014), Forward logistic regression for earth-flow landslide susceptibility assessment in the Platani river basin(southern Sicily, Italy), *Landslides*, Vol. 11, No.4, pp. 639-653.

Guisan, A., Weiss, S.B., and Weiss, A.D. (1999), GLM versus CCA spatial modeling of plant species distribution, *Plant Ecology*, Vol. 143, No.1, pp. 107-122.

Hong, S. (2005), *Binary and Multinomial Logistic Regression*, Kyoyookgwahaksa, Korea.

Hosmer, D.W. and Lemeshow, S. (2000), *Applied Logistic Regression*, Wiley, New York.

Jang, K., Song, J., Park, K., and Chung, J. (2009), An objective procedure to decide the scale factors for applying land-form classification methodology using TPI, *Journal of Korean Forest Society*, Vol. 98, No. 6, pp. 639-645. (in Korean with English abstract)

Lee, H. and Kim, G. (2012), Landslide risk assessment in Inje using logistic regression model, *Korean Journal of Geomatics*, Vol. 30, No. 3, pp. 313-321.(in Korean with English abstract)

Lee, S., Jeong, G., and Park, S.J. (2015), Evaluating geomorphological classification system to predict the occurrence of landslides in mountainous region, *Journal of the Korean Geographical Society*, Vol. 53, No. 4, pp. 485-503. (in Korean with English abstract)

Lee, Y.J., Park, G.A., and Kim, S.J. (2006), Analysis of landslide hazard area using logistic regression analysis and AHP approach, *Journal of the Korean Society of Civil Engineers*, Vol. 26, No. 5, pp. 861-867. (in Korean with English abstract)

Othman, A.A., Gloaguen, R., Andreani, L., and Rahnama, M. (2015), Landslide susceptibility mapping in Mawat area, Kurdistan Region, NE Iraq: A comparison of

different statistical models, *Natural Hazards and Earth System Sciences discuss*, Vol. 3, pp. 1789-1833.

Raftery, A.E. (1995), Bayesian model selection in social research, *Sociological Methodology*, Vol. 25, pp. 111-163.

Weiss, A.D. (2001), Topographic position and landforms analysis, *Annual ESRI International User Conference*, 9-13 July, San Diego, CA, pp. 9-13

Woo, C.S., Lee, C.W., and Jeong Y. (2008), Study on application of topographic position index for prediction of the landslide occurrence, *Journal of Korean Environmental Restoration Technology*, Vol. 11, No. 2, pp. 1-9. (in Korean with English abstract)