

로지스틱 회귀분석을 이용한 인제군 산사태지역의 위험도 평가

Landslide Risk Assessment in Inje Using Logistic Regression Model

이환길¹⁾ · 김기홍²⁾

Lee, Hwangil · Kim, Gihong

Abstract

Korea has been continuously affected by landslides, as 70% of the land is covered by mountains and most of annual rainfall concentrates between June and September. Recently, abrupt climate change affects the increase of landslide occurrence. Gangwon region is especially suffered by landslide damages, because the most of the part is mountainous, steep, and having shallow soil. In this study, a landslide risk assessment model was developed by applying logistic regression to the various data of Duksan-ri, Inje-eup, Inje-gun, Gangwon-do, which has suffered massive landslide triggered by heavy rain in July 2006. The information collected from field investigation and aerial photos right after the landslide of study area were stored in GIS DB for analysis. Slope gradient entered in two ways-as categorical variable and as linear variable. Error matrix for each case was made, and developed model showed the classification accuracy of 81.4% and 81.9%, respectively.

Keywords :landslide, logistic regression, GIS, Inje-gun

초 록

우리나라는 국토의 70%가 산지로 이루어져 있고 연평균 강우량의 대부분이 6월과 9월 사이에 집중되어 산사태로 인한 피해를 지속적으로 입어 왔으며, 최근 급변하는 기후에 따라 그 빈도가 점차 증가하고 있다. 특히, 강원도의 경우 지역적 특성상 대부분 산지로 이루어져 있으며 경사가 가파르고 토심 또한 얇아 산사태에 의해 많은 피해를 입고 있다. 본 논문에서는 2006년 7월 집중호우로 인해 대규모 산사태피해가 발생하였던 강원도 인제군 인제읍 덕산리 지역을 대상으로 로지스틱 회귀분석을 수행하여 산사태 위험도평가모형을 개발하였다. 분석을 위하여 대상 지역의 현장조사 및 피해 직후 촬영된 항공사진을 통해 수집한 정보를 이용하여 GIS DB를 구축하였다. 경사도의 경우 범주형 변수와 연속형 변수로 입력하는 두 가지 방법을 적용하였다. 생성된 예측모형에 대해 점오류분류를 실시한 결과 각각 81.4%와 81.9%의 분류정확도를 보였다.

핵심어: 산사태, 로지스틱 회귀분석, GIS, 인제군

1. 서 론

1.1 연구목적 및 배경

산사태란 사면의 경계에서 발생된 전단파괴의 결과로 흙덩이 또는 돌덩이가 경사면의 하부로 내려오는 것이다

(Skempton & Hutchinson, 1969). 우리나라의 산사태는 주로 집중호우와 태풍이 도래하는 6~9월에 집중적으로 발생하는 특징이 있으며 많은 인명과 재산피해를 유발하고 있다. 더구나 기후 변화에 따라, 1980년대 연평균 231ha였던 산사태 피해 면적은 2000년대 이후 713ha로 3배 이상 증가하였

1) 정희원 · 강릉원주대학교 토목공학과 석사과정(E-mail : killgil21@naver.com)

2) 교신저자 · 정희원 · 강릉원주대학교 토목공학과 부교수(E-mail : ghkim@gwnu.ac.kr)

다. 인명 피해 또한 크게 증가하여 최근 10년간 산사태로 인한 인명 피해가 118명에 달하며, 2011년 한 해에만 43명의 인명피해가 발생하였다. 특히, 2011년의 우먼산 산사태는 도심지에서 발생하여 산태가 산지지역만의 문제가 아니라는 것을 보여주었다.

강원도 지역은 약 81%의 면적이 산지로 형성되어 있다. 해발 100m 이하의 저지대는 강원도 총면적의 5.6%에 불과하며, 50%이상의 면적이 500m이상의 고지대를 이루고 있다. 이러한 산지는 대부분 가파르고 험한 지형이 많고, 유효 토심이 2m 이내로 산사태 발생에 적절한 조건을 가지고 있다. 또한, 동해안과 나란히 길고 높게 발달한 태백산맥이 서남쪽에서 접근해 오는 장마 전선이나 태풍 등의 진로를 막아 집중호우가 빈번히 발생하고 있다(임오빈, 2009). 강원도 지역의 특성상 집중호우로 인해 그림 1과 같은 대규모 재해가 발생되면 산사태로 인한 도로 유실 등으로 신속한 복구지원이 힘든 실정이다. 2002년 '태풍 루사'가 지나갈 때 하루 동안 800mm 이상 쏟아진 폭우로 인하여 발생한 강릉 지역의 산사태와 2006년 평창, 인제 지역의 집중호우로 인한 산사태가 대표적인 예이다. 두 경우 모두 강력한 강우로 유발된 산사태로 인해 많은 인명 및 재산피해가 발생하였는데, 피해 지역의 접근 도로가 유실되어 2차 피해 방지를 위한 복구와 피해 지역 주민 지원이 어려워 많은 문제가 발생되었다. 이와 같은 피해가 반복되는 이유는 산사태가 순식간에 발생할 뿐만 아니라 다른 자연 재해에 비해 예방이 매우 어렵기 때문이다.

산사태 피해를 저감하기 위해서는 위험 지구를 미리 선정하고 관리하는 것이 매우 중요하다. 이러한 위험 지구 선정을 위해서는 산사태가 발생할 지역을 예측하여야 한다. 산사태의 예측을 위해 산사태 유발 인자에 대한 다각도의



그림 1. 2006년 집중호우로 인한 산사태

연구가 진행되어 왔으며, 물리적인 기법보다는 통계적인 기법을 통한 접근이 주로 이루어지고 있다. Dai(2002) 등은 지형, 토지피복, 지질 등의 범주형 자료를 이용한 로지스틱 회귀분석을 통하여 홍콩 Lantau 지역의 산사태 취약성 분석을 실시하였으며, Kawabata(2009) 등은 니가타의 산사태 피해지역에 인공신경망기법을 적용하여 산사태 예측모형을 생성하고 평가하였다. 산림청(1998)은 1980~1983년과 1990~1992년에 전국적으로 발생한 511개소의 산사태 지역에 대해 현지 조사를 실시하여 지형, 임상 등의 자료를 구축하였고, 통계적 방법인 다변량해석법을 이용하여 산사태판정점수표를 제작하였다. 김원영(2004)은 강원도와 경상도 산사태 발생지역에 대해 현장조사를 실시하고 토양 및 지질특성 중심으로 구축된 DB 자료를 로지스틱 회귀분석에 적용하여 산사태 위험도를 산정하였다. 본 논문에서는 로지스틱 회귀분석을 이용하여 산사태 위험도 평가 모형을 생성하고 산사태 위험도를 제작하였다.

1.2 연구대상지역

연구대상 지역은 강원도 인제군 인제읍 덕산리이다(그림 2). 이 지역에서는 2006년 7월 11일에서 20일 사이에 발

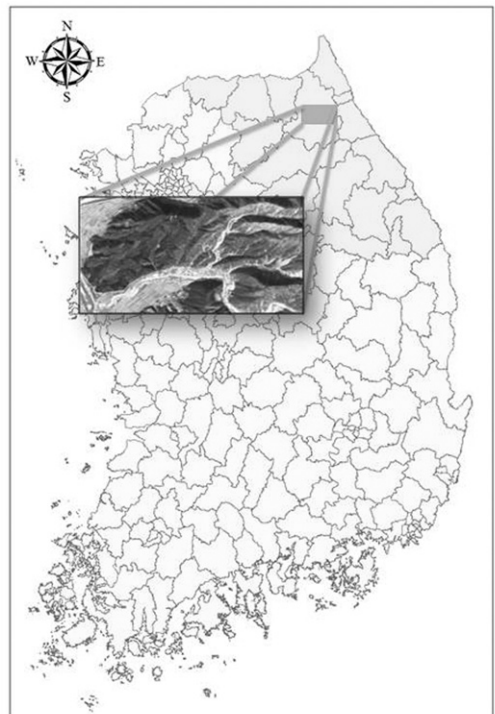


그림 2. 연구대상지역 위치

생한 집중호우로 인해 산사태가 발생하였으며, 7월 15일 오전에 3시간 동안 200mm 내외의 기록적 폭우가 내렸다. 이로 인하여 인제군에만 4,000억 원 이상의 피해가 발생하였고 사면 하부마을에서 주민의 매몰, 사망사고가 발생하였다. 조사 대상 지역의 면적은 8.4km²이며, 최저고도 180m, 최대고도 880m로 매우 가파르고 험준한 지형을 띄고 있다.

2. 연구방법

2.1 자료취득

현장조사와 항공사진 판독을 병행하여 자료를 취득하였다. 산사태 발생 이후 인제읍 지역에 대한 현장조사를 실시하였다(그림3). 레이저 거리 측정기, 클리노미터, 휴대용 GPS 등을 이용하여 산사태 발생부의 위치, 경사, 사면의 형상, 수목의 종류, 수목의 흉고직경 등의 정보를 수집하고 산사태 현장조사야장에 기록하였다(그림4).



그림 3. 현장조사

ID 1		조사 일자 : 2011 년 8 월 4 일									
GPS좌표 (WGS84)						경사 (°)		경사량 (°)			
-N 37.09658		E 128.21441				40		319			
폭(상)	14	폭(중)	18	폭(하)	15	평균값	15.6	깊이	8.8	면적	128.04
깊이(상)	0.8	깊이(중)	1.2	깊이(하)	0.9	평균값	0.97	(m)			
타피 형태						식생종류					
평면	원주	왜기	복합	침엽수	활엽수	대나무	물발	경작지			
특구현황											

그림 4. 현장조사야장

또한 산사태 직후의 항공사진을 이용하여 육안판독에 의한 디지털링 방법으로 산사태 발생지점의 위치정보를 추출하였다. 이 위치정보를 기반으로 수치표고모형(DEM), 임상도(산림청), 토양도(농촌진흥청)에서 산사태 발생 지점 및 부근의 지형, 임상, 토양 정보를 취득하였다. 이러한 방법을 통해 산사태 발생지 92개, 미발생지 107개 지점에 대한 자료를 정리하여 테이블 형식으로 DB를 구축하였다. 이 DB를 통계프로그램에 입력하여 로지스틱 회귀분석을 수행하였다.

2.2 입력 자료

2.2.1 수치표고모형(Digital Elevation Model)

수치표고모형은 표고 자료를 격자 형태로 나타낸 것이다. 본 논문에서는 수치지형도에서 등고선 레이어를 추출하고, 추출된 등고선을 이용하여 해상도 10m급의 수치표고모형을 제작하였다. 제작된 수치표고모형을 바탕으로 GIS툴을 이용하여 경사, 경사향, 사면형상을 제작하였다. 그림 5는 수치표고모형과 산사태 발생 및 미발생 지점을 중첩한 것이다. ▲ 심볼이 산사태 발생지점, × 심볼이 미발생지점이며, 흰색에 가까울수록 높은 고도 값을 가진다.

중심 셀을 기준으로 8방향 경사의 평균을 계산하여 경사도를 제작하였으며 이전에 추출하였던 산사태 발생 및 미발생 지점과 중첩하여 자료를 취득하였다. 경사는 연속형 변수와 범주형 변수 두 가지로 정리하였다. 연속형 변수로는 원래의 경사가 아닌 산사태 발생지의 평균 경사와 해당 지점 경사의 차이값을 사용하였다. 식 (1)은 경사인자 계산식이며 발생지역의 평균 경사는 28.44°이다.

$$|산사태 발생지역 평균 경사 - 각 셀의 경사| \quad (1)$$

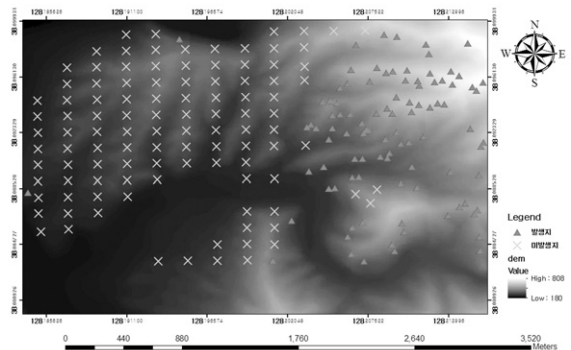


그림 5. 수치표고모형 중첩

범주형 변수는 5° 미만, 5°-10°, 10°-15°, 15°-20°, 20°-25°, 25°-30°, 30°-35°, 35°-40°, 40° 이상의 총 9등급으로 분류하였다. 그림 6은 경사도와 산사태 발생 및 미발생 지점을 중첩한 것이다. 흰색에 가까울수록 높은 경사도 값을 가진다.

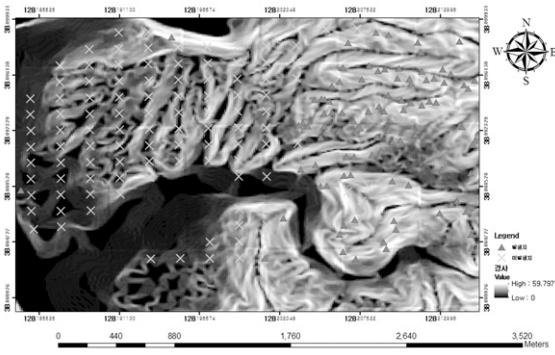


그림 6. 경사도 중첩

경사향은 경사도와 마찬가지로 중심 셀을 기준으로 경사를 이용하여 제작하였다. 계산된 경사향은 범주형 변수로 변환하였으며, F, N, NE, E, SE, S, SW, W, NW의 총 9등급으로 재분류하였다.

사면형상은 중심 셀을 기준으로 4방향의 경사를 참고하여 판단한다. 사면형상의 경우 현장에서 수집한 자료는 파괴 후의 형상이므로 파괴 전의 형상을 추정하기 위해 현장 조사 자료를 사용하지 않고 DEM 자료만을 사용하였다. 사면형상은 오목, 평면, 볼록 3가지로 분류하였다.

2.2.2 임상도

임상에 대한 정보는 산림청에서 제공하는 1:25,000 수치 임상도를 이용하여 수집하였다. 임상도의 주요 속성정보에는 수종, 영급, 경급 밀도가 있다. 각각의 속성정보를 분류하여 레스터자료로 변환하였고, 산사태 발생 및 미발생 지점의 위치정보와 중첩하여 자료를 취득하였다. 임상도의 자료는 범주형 변수이므로 재분류 과정을 거치지 않고 분석에 사용하였다. 그림 7은 임상도(영급)와 산사태 발생 및 미발생 지점을 중첩한 것이다. 보라색이 비산림, 연두색이 2영급, 붉은색이 3영급, 흰색이 4영급, 녹색이 5영급으로 높은 등급을 가지는 지역에 산사태 발생지가 많이 분포하는 것을 알 수 있다.

수종은 약 20여 가지로 분류되나 연구대상지역에 분포하는 수종은 소나무림, 소나무인공림, 낙엽송림, 제지였다.

영급은 수목의 나이를 등급화한 것이다. 대상지역에는 2영급에서 5영급사이의 수목이 분포하였다. 경급은 수목의 고직경을 등급화한 것인데, 대상지역에는 소경목과 중경목이 주로 분포하였다.

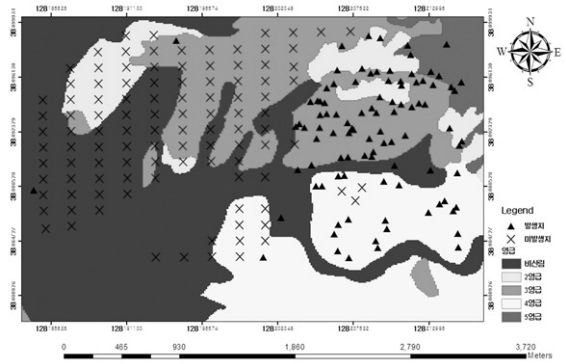


그림 7. 임상도(영급) 중첩

2.2.3 토양도

토양에 대한 정보는 농촌진흥청에서 제공되는 1:25,000 수치 토양도를 이용하여 수집하였다. 토양도의 여러 속성정보 중 배수등급, 심토자갈함량, 표토토성이 분석에 활용되었다. 토양도 자료는 임상도자료와 같이 범주형 변수이므로 재분류 과정을 거치지 않았다. 그림 8은 토양도(심토자갈함량)와 산사태 발생 및 미발생 지점을 중첩한 것이다. 파란색이 <10%, 회색이 10~35%, 연두색이 >35%로 연구대상지역에는 주로 10% 이하의 지역이 가장 많이 분포했다.

배수등급은 토양을 그 배수 능력에 따라 5등급으로 분류한 것이다. 대상지역의 배수등급은 매우양호, 양호, 약간양호로 분류되어 있다. 심토자갈함량은 심토의 직경 2mm이

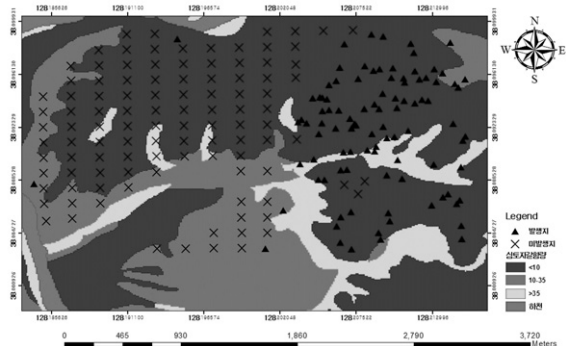


그림 8. 토양도(심토자갈함량) 중첩

상 자갈의 함량을 나타내는 것이다. 10% 미만, 10-35%, 35% 초과로 구분되며 대상지역에 모두 분포되어 있었다. 표토 토성은 표토를 미사, 모래, 점토의 함량에 따라 구분한 것이다. 대상지역은 사양토와 양토로 구성되어 있다.

3. 산사태 위험도 평가

3.1 로지스틱 회귀분석

3.1.1 로지스틱 회귀분석 이론

회귀분석이란 둘 또는 그이상의 변수들 간의 관계를 파악함으로써 종속변수의 값을 독립변수들로 부터 설명하고 예측하는 통계적 기법이다. 또한 회귀분석은 종속변수와 독립변수들 간의 관련성을 분석하는데 있어 매우 강력한 분석력을 갖고 있으며 적응성이 뛰어난 특성을 가지고 있다. 일반적인 회귀분석식은 식 (2)와 같다.

$$E(Y|X) = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (2)$$

여기서 b_0 는 절편이며 b_1 에서 b_n 까지의 계수는 각 독립변수에 대한 회귀 계수이다. $E(Y|X)$ 는 각 독립변수 X 가 특정 값을 취할 때 Y 의 기댓값 또는 예측값이 된다. $E(Y|X)$ 는 각 독립변수의 값에 따라 어떤 값이라도 취할 수 있으므로 이론적으로 $E(Y|X)$ 의 범위는 $-\infty$ 에서 ∞ 가 된다. 그러나 종속변수가 산사태의 발생, 미발생과 같은 이항형 변수일 때 선형회귀분석을 적용하면 독립변수와 종속변수의 관계를 제대로 설명하지 못하며, 독립변수가 선형식으로 표현되기 보다는 S자 형태를 따르는 곡선에 가깝게 된다. 이러한 독립변수와 종속변수의 관계를 설명하는 함수가 로지스틱(logistic)함수이다. 이 함수를 사용함으로써 확률의 개념을 지니고 있는 종속변수 값이 0과 1사이를 벗어나지 않도록 할 수 있다. 즉, 독립변수의 값이 아무리 커도 예측된 확률 값은 0이나 1에 근접하지만 0보다 작거나 1을 넘어서지 않는다. 로지스틱 함수를 수식으로 표현하면 식 (3)과 같다.

$$E(Y|X) = \frac{\exp(b_0 + b_1X)}{1 + \exp(b_0 + b_1X)} \quad (3)$$

독립변수 X 를 이용하여 예측한 Y 값은 $E(Y|X)$ 이며 종속변수 Y 가 이항형변수일 경우 $E(Y|X)$ 는 확률 개념을 가진다. 즉, 종속변수 Y 의 예측된 확률 값은 독립변수 X 값에 달려 있게 되는 것이다(홍세희, 2005).

위에서 언급하였던 바와 같이 산사태의 경우 종속변수가 발생(1)과 미발생(0)으로 구분되는 이항형 변수 이므로 일반회귀 모형을 그대로 적용하기 하기 어려우나 로지스틱 함수를 이용하여 용이 하게 분석할 수 있다.

3.1.2 로지스틱 모형의 추정

일반적인 회귀 분석에서는 미지수인 절편과 회귀계수를 최소 자승화 기준을 가장 만족 시키는 값으로 추정한다. 그러나 종속변수가 연속 변수가 아닌 이항형 변수인 경우에는 최소 자승화 기준을 이용한 추정방법대신 최대우도법(maximum likelihood)이라는 추정 방법을 적용한다. 최대우도법에서는 관찰 자료의 가능성을 최대화 하는 값으로 미지수의 추정치를 구한다. 최대우도법을 적용하기 위해서는 먼저 우도함수(L)를 설정하고 이 우도함수를 최대화 하는 값으로 미지수의 추정치를 결정한다. 여기서 우도함수란 미지수에 따른 관찰 자료의 가능성을 함수로 표현한 것을 말한다. 우도함수 L 은 식 (4)와 같이 나타낼 수 있다(홍세희, 2005).

$$L = \prod [p_i^{y_i} [1 - p_i]^{1 - y_i}] \quad (4)$$

식 (4)에서 y_i 는 산사태 발생, 미발생의 이항형 값을, p_i 는 로지스틱 모형을 통해 예측된 확률을 의미한다.

3.1.3 로지스틱 회귀모형 및 회귀계수에 대한 검증

로지스틱 회귀모형의 적합도를 평가하기 위해 분석모형(M)과 포화모형(S)을 비교한다. 두 모형을 비교하는 식은 식 (5)와 같다.

$$-2\ln L_M - (-2\ln L_S) \quad (5)$$

여기서 $\ln L_M$ 은 분석모형의 로그 우도 함수를, $\ln L_S$ 는 포화 모형의 로그 우도 함수를 나타낸다. 모형의 적합도가 좋아질수록 $-2\ln L$ 은 0에 가까워지며 모형의 적합도가 나빠질수록 양수로 커진다. 로그연산을 이용하여 식 (5)를 식 (6)과 같이 변형할 수 있다.

$$-2\ln L \left[\frac{L_M}{L_S} \right] \quad (6)$$

여기서 $\frac{L_M}{L_S}$ 를 우도비라고 하며, 여기서 얻은 통계값을 이용한 검증을 우도비 검증이라고 한다. 로지스틱 회귀분석에서 각 회귀 계수가 통계적으로 유의한지를 검증하는 방

법은 두 가지가 있다. Wald 검증을 이용하면 해당 로지스틱 회귀계수가 자유도 1의 χ^2 분포에서 유의미한지 알아볼 수 있다. 우도비 검증을 이용하면 해당 변수를 포함한 모형과 포함하지 않은 모형의 $-2\ln L$ 차이를 구하여 이 차이가 자유도 1인 χ^2 분포에서 유의미한지 알아볼 수 있다. 두 방법 모두 장단점이 있으나 Wald 검증 보다는 우도비 검증의 검증력이 높으므로 우도비 검증을 많이 사용하고 있다(홍세희, 2005).

3.2 산사태 위험도 평가모형 개발

로지스틱 회귀분석에 앞서 독립변수가 범주형인 경우 더미코딩(dummy coding)을 할 필요가 있다. 더미코딩이란 범주형 변수가 독립변수로 사용될 경우 등간척도 이상이라는 선형성 조건에 적합하지 않은 문제를 해결하기 위해 고안된 방법으로 측정된 변수 값들을 구분할 수 있도록 하기 위해 가변수를 생성하고 자신의 집단은 1 나머지 집단은 0으로 변수로 코딩하는 것이다. n개의 집단을 더미 코딩하는 경우 n-1개의 더미변수가 필요하다. 마지막 더미변수는 앞의 더미변수 코딩에 따라 자동적으로 결정되어 독립적인 변수로 볼 수 없기 때문에 더미변수로 활용하지 않고 기준집단으로 활용한다. 기준집단은 다른 더미변수들과의 관계를 설명할 때 사용한다. 단 더미변수들 간의 관계에 대한 설명은 회귀식으로 확인할 수 없다(홍세희, 2005). 본 연구에서는 더미코딩의 마지막 변수를 기준집단으로 활

용하여 로지스틱 회귀분석을 수행하였다. 최종적으로 분석 수행된 자료는 경사, 수종, 영급, 배수등급, 심토자갈함량, 표토토성이다. 나머지 경사향, 사면형상, 경급은 통계학적으로 유의하지 않은 것으로 나타나 분석에서 제외하였다. 로지스틱 회귀분석의 결과는 표 1과 같다. 경사도는 35°-40° 인 경우 산사태에 가장 취약한 것으로 나타났는데, 이는 Dai(2002)의 연구결과와 비슷하다. 수종은 소나무 인공림과 낙엽송이 산사태와 가장 높은 연관성을 보였으며, 영급은 임목의 수령이 높아질수록 산사태 발생 위험이 컸다. 또한 배수가 잘될수록, 심토자갈함량이 35% 이상일 때, 표토토성이 사양토보다는 양토일수록 산사태에 취약했다.

로지스틱회귀분석의 결과는 0에서 1사이에 분포하기 때문에 산출된 값을 확률값으로 판단할 수 있다. 따라서 산사태 발생의 판정 기준을 0.5로 하여 0.5보다 크면 발생으로 0.5보다 작으면 미발생으로 분류하는 것이 일반적이다. 정오분류 결과 산사태 발생지역 92개중 82개를 옳게 예측하여 89.1%의 분류정확도를 보였으며, 산사태가 발생하지 않은 지역은 107개 중 80개를 옳게 예측하여 74.8%의 분류정확도를 보였다. 산사태 발생과 미발생 지역을 종합하면 약 81.4%의 분류정확도를 보였다. 이는 Dai(2003)의 산사태 위험도 예측모형의 분류정확도 84.9%와 비슷한 결과이며, 이용준(2006)이 제시한 로지스틱 회귀모형의 분류정확도 76.8%보다는 다소 높은 값이다.

표 1. 로지스틱 회귀분석 결과(범주형경사)

변 수	범 주	계 수	변 수	범 주	계 수		
상수		-5.244	영급	2영급	22.125		
	경사(°)	<5		-1.885	3영급	43.185	
		5-10		0.767	4영급	44.652	
		10-15		0.215	5영급	63.603	
		15-20		0.399	비산림		
		20-25		1.482	배수등급	매우양호	36.539
		25-30		1.577		양호	21.583
		30-35		0.792		약간양호	
		35-40		1.642		<10	-16.724
		수종		≥40		10-35	-19.688
소나무림	-42.127		>35				
낙엽송림	-0.620		심토자갈함량 (%)				
제지	-20.093			표토토성	사양토	-16.386	
	비산림		양토				

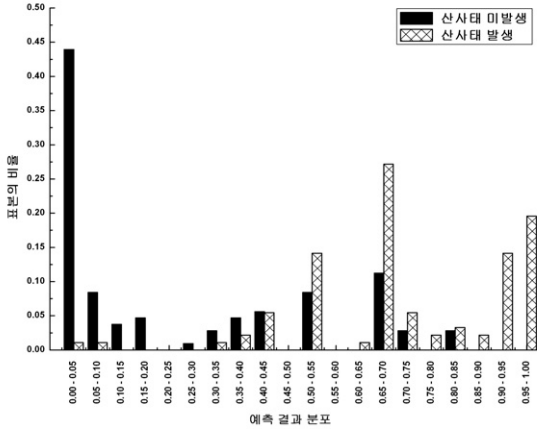


그림 9. 산사태 위험도 예측 결과 분포도(범주형 경사)

산사태위험도 예측모형의 확률 분포를 보기 위해 로지스틱 회귀분석에 적용한 각각의 DB를 산사태위험도 예측 모형에 적용하였다. 산정된 예측 결과를 그림 9와 같이 나타내었다. 산사태 발생지역은 대부분 0.65이상의 값을 보였으며, 발생하지 않은 지역은 대부분 0.2이하의 분포를 보였다. 분류기준 0.5의 근사값인 0.40 - 0.55사이의 값이 분포하여 통계적으로 유의한 인자를 추가하여야 더 정확한 모형을 개발할 수 있을 것으로 사료된다.

생성된 로지스틱 회귀식을 이용하여 산사태위험도평가 지도를 제작하였다. 각각 인자를 래스터 자료로 변환하고, GIS 도구를 이용하여 회귀식의 계수로 등급화 과정을 거침으로서 연구대상지역에 대한 산사태위험도평가지도를 제작하였다. 붉은색에 가까울수록 산사태 발생에 높은 확률

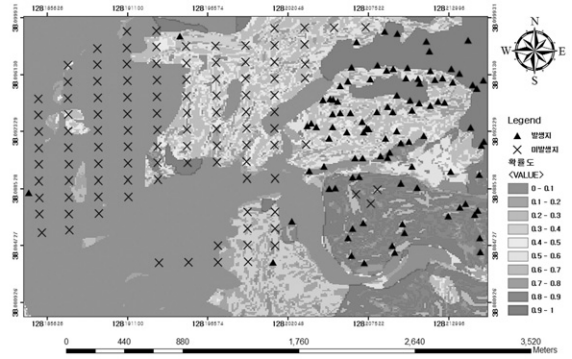


그림 10. 산사태위험도평가지도(범주형 경사)

을 보이며, 녹색에 가까울수록 낮은 확률을 보인다. 단, 연구대상지역에 인자는 존재하나 로지스틱 회귀식 표본에 포함되어있지 않은 인자로 인하여 확률이 잘못 산정되는 경우도 발생하였다. 이는 추후 연구대상지 확대를 통해 해결 될 것으로 사료된다.

경사도를 범주형 변수가 아닌 산사태 평균 경사와의 차이값인 선형 변수로 입력하여 로지스틱 회귀분석을 수행한 결과는 표 2와 같다. 전체적으로 표 1의 회귀분석 수행 결과와 매우 비슷한 결과를 보였다.

표 2의 계수를 로지스틱 회귀식에 적용하여 각각의 자료에 대해 정오분류를 실시한 결과 산사태 발생지역 92개중 81개를 발생으로 분류하여 88%의 분류정확도를, 산사태가 발생하지 않은 103 지역 중 82개를 미발생으로 분류하여 76.6%의 분류정확도를 보였으며, 이를 종합한 총 분류정확

표 2. 로지스틱 회귀분석 결과(산사태 평균 경사와의 차이)

변 수	범 주	계 수	변 수	범 주	계 수
상수		-3.401	영 급	2영급	22.088
산사태 평균 경사와의 차이(°)		-0.840		3영급	43.196
수 종	소나무림	-42.409		4영급	44.466
	낙엽송림	-0.141		5영급	63.328
	제지	-19.913		비산림	
심토자갈함량 (%)	비산림		배수등급	매우양호	36.240
	<10	-16.623		양호	21.144
	10-35	-19.332	표토토성	약간양호	
>35		사양토		-16.495	
				양토	

도는 81.9%이다. 경사도를 범주형 변수로 두고 수행한 결과에 비해 약 0.5% 개선된 결과를 얻을 수 있었다.

그림 11은 그림 9와 같이 각각의 로지스틱 회귀분석에 적용한 DB를 대상으로 예측한 분포도이다. 실제 산사태가 발생하지 않은 지역의 경우 0~0.05사이에서 45%이상이 분포하였고, 산사태가 발생한 지역은 0.5이상에 고르게 분포하였다. 범주형 경사를 이용한 로지스틱 회귀식 결과분포(그림 6)에서의 0.5~0.55사이의 분포가 줄어들고 0.45~0.5, 0.55~0.6의 분포가 증가 하여 분류정확도 기준인 0.5 이내의 값은 크게 변하지 않은 것으로 사료된다. 이전 회귀식과 마찬가지로 산사태 발생에 통계적으로 유의한 인자가 추가된다면 모형의 정확도를 향상시킬 수 있을 것으로 생각된다.

그림 12는 산사태 평균 경사와의 차이를 이용한 로지스틱 회귀식을 바탕으로 산사태위험도 평가지도도를 제작한 것이다. 범주형 경사를 이용한 산사태위험도 평가지도보다 전체적으로 확률이 낮게 분포하였다.

4. 결론

본 연구에서는 2006년 집중호우로 인한 산사태로 대규모 피해가 발생한 강원도 인제군 인제읍 덕산리 지역을 대상으로 산사태 위험도 평가모형을 제작하였다. 산사태 피해지역에 대한 현장조사를 통하여 산사태 발생부의 위치, 경사, 수종 등 각종정보를 취득하였다. 현장조사를 통해 수집하기 어려운 자료를 확보하기 위해 산사태 직후 촬영된 항공사진을 이용하였으며, 육안판독에 의한 디지털링 방법으로 산사태 발생지점 92개소와 미발생지점 107개소의 위치를 추출하였다. 분석에 필요한 각 지점의 자료는 추출된 위치정보를 경사도, 경사향도, 사면형상도, 임상도(수종, 경급, 영급), 토양도(배수등급, 표토토성, 심토자갈함량)와 중첩분석하여 취득하고, DB에 저장하였다. 구축된 DB를 이용하여 로지스틱 회귀분석을 수행하였으며, 이 과정에서 경사향, 사면형상, 경급은 통계적으로 유의하지 않아 분석에서 제외되었다. 최종적으로 경사도, 수종, 경급, 영급, 배수등급, 표토토성, 심토자갈함량이 고려된 로지스틱 회귀식을 생성하였다. 로지스틱 회귀식은 경사도를 범주형 변수로 입력하는 방법과, 선형(산사태 평균 경사와의 차이값) 변수로 입력하는 방법으로 수행하였다. 생성된 회귀식을 DB에 적용하여 정오분류를 실시한 결과 각각 81.4%와 81.9%의 분류 정확도를 보였다. 분류결과, 분류기준 0.5의 근사값인 0.4~0.6사이의 값이 많이 분포하였으며, 통계적으로 유의한 인자를 추가하여야 더욱더 정확한 모형을 생성할 수 있을 것으로 사료된다. 분류정확도 오차의 원인은 산사태 발생 후 시간경과에 따른 조사의 정밀성 부족과 로지스틱 회귀분석에 사용된 표본수의 부족, 각각의 기관에서 제작한 GIS자료와 실제 지형, 임상, 토양 특성과의 오차 때문인 것으로 추정된다. 추후 연구대상지를 확대하여 로지스틱 회귀분석에 적용할 표본수를 늘리고 대축척 GIS 자료를 이용하여 자료의 정밀도를 높인다면 더욱 개선된 산사태위험도 평가모형을 개발할 수 있을 것이다.

감사의 글

이 논문 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 연구임 (No. 2010-0006844)

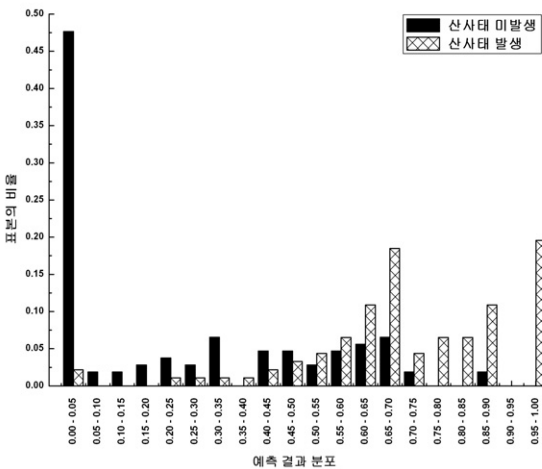


그림 11. 산사태 위험도 예측 결과 분포도(산사태 평균 경사와의 차이)

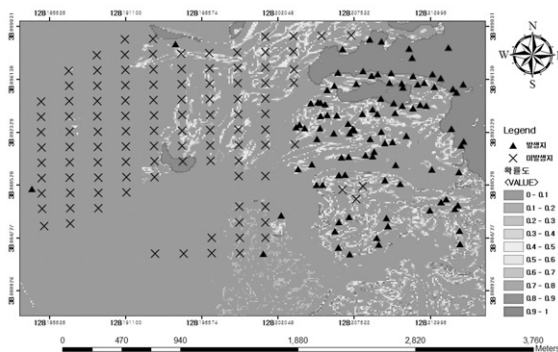


그림 12. 산사태위험도평가지도(산사태 평균 경사와의 차이)

참고문헌

- 김원영 (2004), 산사태 위험도 산정시스템 및 피해저감기술 개발, 과학기술부 자연재해방재기술개발사업, 한국지질자원연구원, pp. 3-8.
- 산림청 (1998), 사방기술교본, pp. 416-417.
- 이용준, 박근애, 김성준 (2006), 로지스틱 회귀분석 및 AHP 기법을 이용한 산사태 위험지역분석, 대한토목학회지, 대한토목학회, 제 26권, 제 5호, pp. 861-867.
- 임오빈 (2009), 호우 피해지 산사태 발생의 특성 및 복구에 관한 연구 - 홍천지역을 중심으로, 석사학위논문, 강원대학교, pp. 13-14.
- 홍세희 (2005), 이항 및 다항 로지스틱 회귀분석, 교육과학사, pp. 1-141.
- 황제선, 박초롱, 윤찬영, 김기홍 (2011), 토석류 현장조사기법의 표준화, 2011한국지형공간정보 학회춘계학술대회 논문집, 한국지형공간정보학회, pp. 215-216.
- Kawabata, D. and Bandibas, J. (2009), Landslide susceptibility mapping using geological data, a DEM from ASTER images and an artificial Neural Network(ANN), *Geomorphology*, Vol. 113, No. 1, pp. 97-109.
- Dai, F. C. and Lee, C. F. (2002), Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong, *Geomorphology*, Vol. 42, No. 3, pp. 213-228.
- Dai, F. C. and Lee, C. F. (2003), A spatiotemporal probabilistic modelling of storm-induced shallow landsliding using aerial photographs and logistic regression, *Earth Surface Processes and Landforms*, Vol. 28, No. 5, pp. 527-545.
- Skempton, A. W. and Hutchinson, J. N. (1969), Stability of natural slopes and embankment foundation, *Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering*, Vol. 3, pp. 291-340.

(접수일 2012. 05. 29, 심사일 2012. 06. 08, 심사완료일 2012. 06. 26)