

예측비율곡선을 이용한 GIS 기반 산사태 예측 모델의 정량적 비교

지광훈* · 신진수** · 박노옥*

한국지질자원연구원 국가지질자원정보센터*, 중앙항공 지리정보연구소**

Quantitative Analysis of GIS-based Landslide Prediction Models Using Prediction Rate Curve

Kwang-Hoon Chi*, Jin-Soo Shin**, and No-Wook Park*

National Geoscience Information Center, KIGAM*, Chung-Ang Aerosurvey Co., Ltd.**

Abstract : The purpose of this study is to compare the landslide prediction models quantitatively using prediction rate curve. A case study from the Jangheung area was used to illustrate the methodologies. The landslide locations were detected from remote sensing data and field survey, and geospatial information related to landslide occurrences were built as a spatial database in GIS. As prediction models, joint conditional probability model and certainty factor model were applied. For cross-validation approach, landslide locations were partitioned into two groups randomly. One group was used to construct prediction models, and the other group was used to validate prediction results. From the cross-validation analysis, it is possible to compare two models to each other in this study area. It is expected that these approaches will be used effectively to compare other prediction models and to analyze the causal factors in prediction models.

Key Words : Prediction Rate Curve, Landslide, Prediction Model

요약 : 본 연구는 예측비율곡선을 이용한 산사태 예측 모델의 정량적 비교분석을 목적으로 하여 수행되었다. 1998년 8월 집중호우로 산사태가 발생한 경기도 장흥지역을 대상으로 위성영상과 현장 답사를 통하여 산사태 발생위치를 확인하였고, GIS 기반의 다양한 산사태 관련 공간 정보를 구축하였다. 사용된 공간통합 방법은 결합 조건부 확률과 certainty factor 이며, 산사태 발생 지역을 무작위로 2개의 그룹으로 나누어서 한 개의 그룹은 예측도 작성에 사용하였으며, 나머지 그룹은 예측 결과의 검증에 사용하였다. 예측비율곡선 작성을 통해 두 예측모델의 정량적 비교가 가능하였다. 본 연구에서 사용된 예측비율곡선 작성은 추후 다른 모델의 정량적 비교 분석 및 오차 분석을 하는 데에도 사용될 수 있을 것으로 기대된다.

1. 서론

산사태는 폭우나 또는 지진에 의해 일시에 산

의 일부가 무너져 내려오는 것으로(김상규, 1988), 우리나라에서는 매년 산사태로 인해 많은 인명 및 재산피해를 내고 있으며, 피해 복구에 많은 시간

2001년 5월 15일 접수 : 2001년 8월 25일 수리.

과 인력을 들이는 등 상당한 경제적인 손실이 발생하고 있다. 전국토의 70%가 산지로 이루어져 있는 우리나라는 강우가 주로 여름에 집중되기 때문에, 산사태로 인한 피해가 이시기에 집중적으로 발생되고 있다(홍원표, 1990).

이와 같은 산사태로 인한 인명피해와 재산피해를 예방하거나 줄이기 위하여 산사태의 과학적인 분석기술 개발이 필요한 실정이다. 하지만 이러한 산사태를 분석하기 위해 필요한 각종 공간정보를 현장조사에 의해 획득하기 위해서는 많은 시간, 인력, 비용을 필요로 하게 된다. 반면, 원격탐사와 GIS 기술을 이용하게 되면, 종래보다 현장조사에 필요한 시간, 인력 및 비용을 크게 줄일 수 있으며, 대상 지역에 대한 공간적인 예측 모델링이 가능하다.

이러한 배경에서 그동안 산사태 취약성 분석을 위한 연구가 많이 이루어져 왔다. 국내의 산사태 연구는 암반공학적인 접근으로, 공간자료의 속성 값에 대하여 가중치를 부여하여 통합하는 연구가 대부분이었다(김윤중 외, 1991; 김원영 외, 1998). 최근에 와서야 GIS를 기반으로한 산사태 취약성 분석은 이루어지고 있으며, DB 구축이 이루어지고 있다(아사로, 2000). 또한 대부분의 연구의 결과는 통계적 유의미성을 갖는 예측(prediction)이라기 보다는 기존 산사태 발생 자료와 공간 정보의 연관성 분석을 통한 적합성을 분석하는 추정(assessment)의 제시였다.

외국의 경우, Chung and Fabbri(1993, 1998)에 의해 조건부 확률, certainty factor, 퍼지 연산 등의 통합방법을 이용하여 다양한 지역에 대한 객관적이고 정량적인 산사태 예측도의 작성이 시도된바 있다. 또한 통합 결과의 타당성 검증을 위하여 시간적으로 구분되는 두 그룹의 산사태 발생 위치를 이용하여 여러 가지 확률적인 통합방법들을 비교하였다(Chung and Fabbri, 1999). 이렇게 작성된 예측도의 정량화 검증은 일반적으로 시간과 공간에 대해서 적용할 수 있다(Chung and Fabbri, 2001). 시간의 경우, 산사태 발생자료를 두 개의 시간 그룹, 예를 들어 과거 발생 자료와 미래 발

생자료로 나누어서 과거 발생자료를 기반으로 예측 모델을 작성하고 예측 결과가 얼마나 미래 발생 자료를 잘 설명할 수 있는지를 검증하는 것이다. Chung and Fabbri(1999)는 Colombia Rio Chincina 지역을 대상으로 1960년 이전에 발생한 산사태 발생자료를 가지고 작성한 예측 모델로부터 1960년 이후에 발생한 산사태 발생자료를 설명하였다. 공간의 경우, 연구지역을 몇 개의 영역으로 나누고 각각의 영역에 대해서 예측한 결과를 모자이크하여 최종 예측 모델을 작성하는 방법과 산사태 발생자료를 임의(random)로 나누어서 특정 자료는 예측 모델 작성에 사용하고, 모델 작성에 사용하지 않은 자료를 가지고 예측 능력을 비교하게 된다.

본 연구에서는 산사태 취약성 분석을 위하여 원격탐사 자료와 공간정보를 이용한 확률 모델을 제시하고, 예측비율곡선(prediction rate curve)을 이용하여 예측도의 공간적 추정 능력의 정량적 비교를 수행하고자 하였다. 경기도 장흥지역을 대상으로, 우선 산사태 분석에 중요한 산사태 발생 위치를 다양한 지구관측 위성 자료 및 현지답사로부터 추출하였으며, GIS 기반의 공간통합방법으로 결합 조건부 확률(joint conditional probability model)과 certainty factor를 이용하여 산사태 예측도를 작성하였다. 그리고 예측비율곡선(prediction rate curve)을 작성하여 두 통합방법의 예측능력을 정량적으로 비교하였다.

2. 사례연구 지역 및 공간 정보 구축

연구 지역은 행정구역상 경기도 양주군 백석면 남부와 장흥면 북부, 의정부시 서부와 주내면 일부를 포함하는 지역이다. 지리적 좌표상으로는 동경 126°56' ~ 127°01', 북위 37°43' ~ 37°46' 사이에 위치하고 있다(Fig. 1). 이 지역은 1998년 8월 4일부터 8월 9일 사이에 내린 집중호우에 의해 많은 수의 산사태가 발생하였다. 특히, 이 시기의 집중호우로 인한 산사태는 경기도에서만 166명의 실종

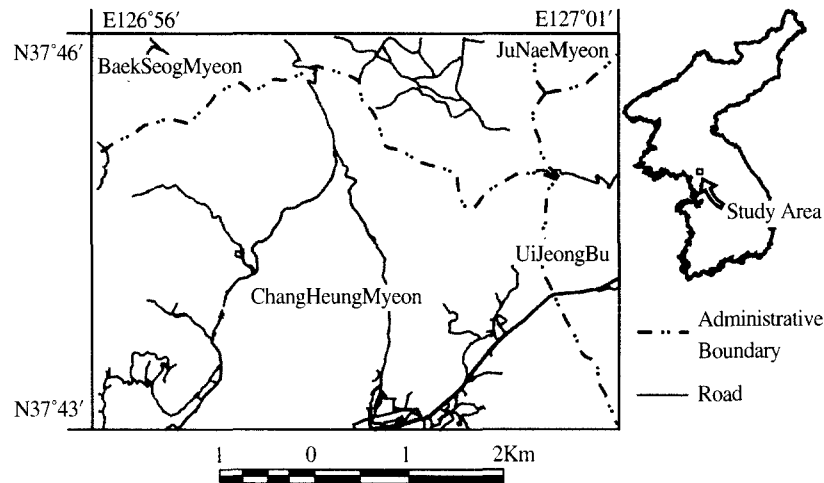


Fig. 1. Location map of the study area.

Table 1. GIS data layers used in this study.

Data Layers	
Landslide location data	Soil map**
Slope map	Lineament density map
Aspect map	Distance from drainage pattern
Forest data sets*	Distance from paved roads
	Distance from unpaved roads
	Geological map***

* Forest map of Dukjung, Kwangjuk, Koyang, Uijeongbu, 1:25,000, Korea forest service, Korea forest research institute, 1992

** Reconnaissance soil map of Korea, Kyungki province(I), 1:25,000, Institute of plant environment, office of rural development, 1976

*** Geological map of Seoul-Namcheonjum, 1:50,000, Korea institute of geology, mining and materials, 1999

자와 사망자가 발생하는 인명피해를 일으켰으며, 수많은 가옥이 완파되고 도로, 비닐 하우스, 묘지 등이 유실되는 재산피해를 발생시켰다.

공간 통합에 사용된 자료로 의정부, 고양, 광적, 덕정 1:25,000 도폭의 수치지도를 이용하여 수치표고모델(DEM)을 작성하였고, 이로부터 사면경사도와 사면방위도를 작성하였다. 수치지형도로부터 포장 도로, 비포장 도로 및 수계망을 추출한 뒤 버퍼링을 수행하여 세 가지 종류의 주제도를 작성하였다. 임상자료로는 영급(wood age), 경급(wood diameter), 밀도(wood density), 수목(wood type)을 작성하였으며, 간이산림토양도를 수치화하여 토양

도를 작성하였다. 또한 선구조와의 관련성을 알아보기 위하여 수치지형도로부터 추출한 선구조 빈도수와 길이를 결합한 선구조 밀도도를 작성하였다. 이밖에 1:250,000 축척의 지질도를 사용하여 주제도를 만들었다(Fig. 2, Table 1).

그리고, 구축된 주제도의 통합에 기준이 되며, 통합 후 교차 검증에 사용될 산사태 발생위치는 지구관측위성영상과 현장조사를 바탕으로 표시하였다. 가용할 수 있는 연구지역의 SPOT-2 Panchromatic 영상, 다목적실용위성의 EOC 영상 및 IRS-1C 영상을 비교하였다. EOC 영상은 2000년 3월 9일 자료, SPOT Panchromatic은 산사태가

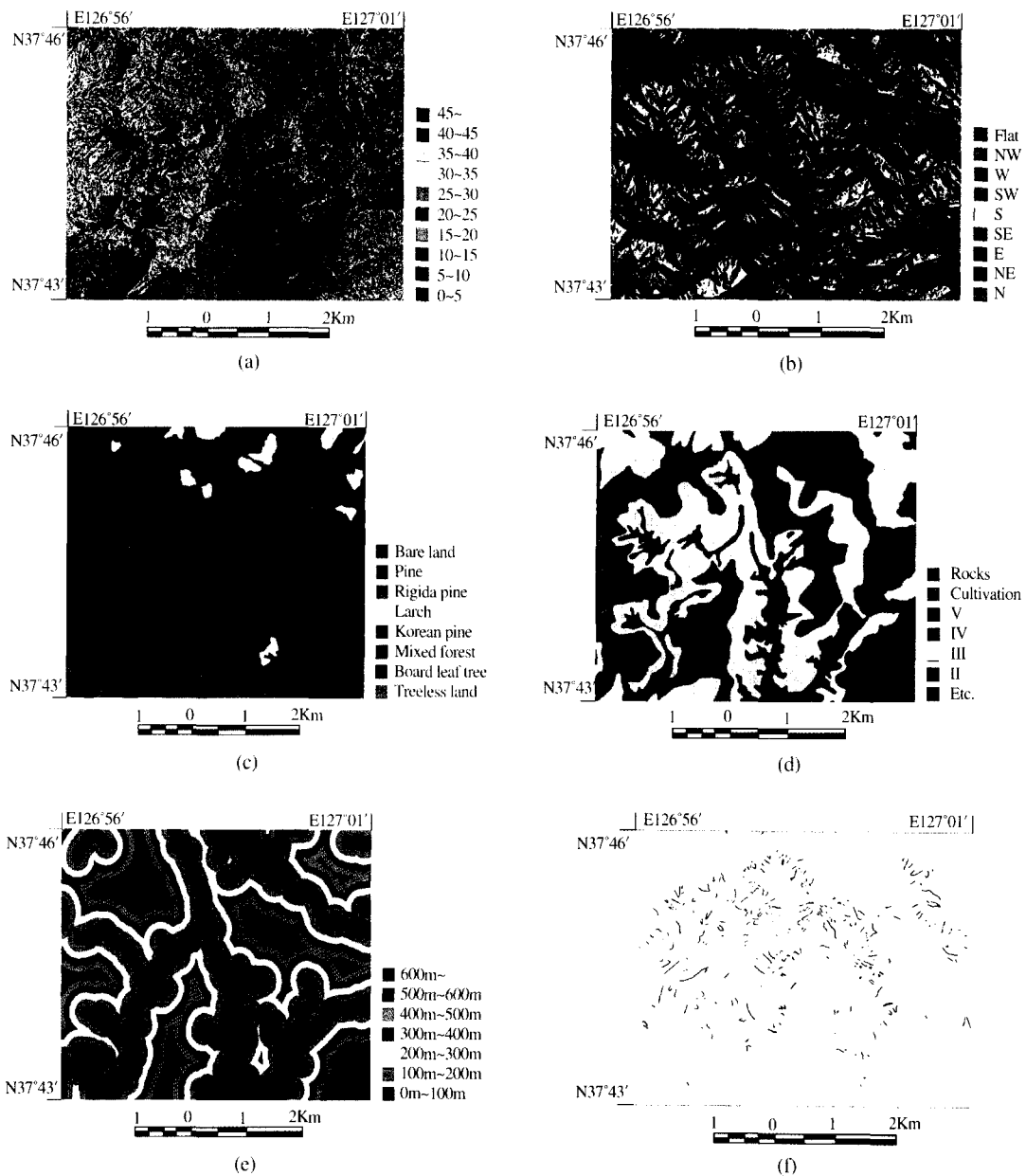


Fig. 2. Spatial data sets in Jangheung area, (a) slope, (b) aspect, (c) forest type, (d) soil map, (e) distance from drainage pattern, (f) landslide location map.

발생한 직후인 1998년 11월 3일에 촬영된 자료, IRS-1C영상은 1999년 1월 26일에 촬영된 영상을 사용하였다(Fig. 3). 각 위성별로 산사태 발생지역의 탐지를 정성적으로 비교분석한 후, 현장조사를 통해 산사태 발생위치를 표시하였다. 이러한 과정

으로 표시되어진 산사태는 주로 지형적인 특성과 인위적인 절개면을 따라 발생하였으며, 발생위치는 총 359개이다. 탐지된 산사태 위치 중에서 유동에 의한 피해 입은 지역은 산사태 발생에 대한 정보를 제공하지 않으므로 제거하였다.

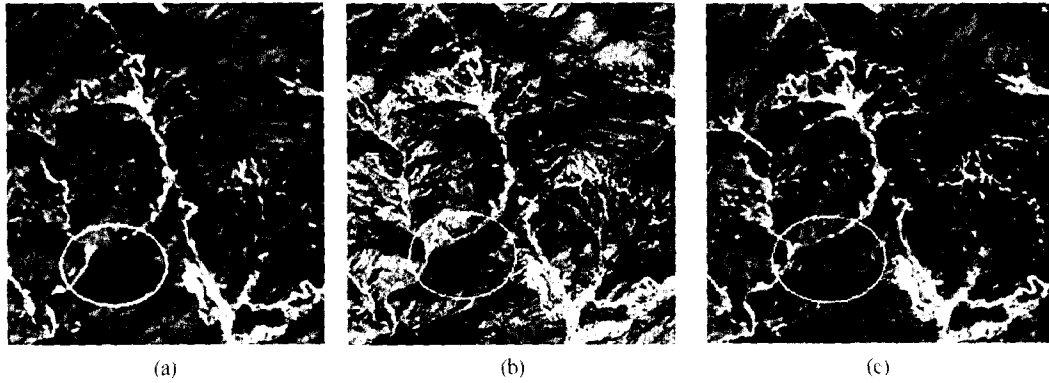


Fig. 3. Comparison of remote sensing data for detecting landslide locations.

(a) Spot (1998. 11. 3) Azimuth 169.4° Elevation 36.7°, (b) IRS 1-C(1999. 1. 26) Azimuth 153.3° Elevation 43.4°, (c) Komsat EOC (2000. 3. 9) Azimuth 147.9° Elevation 42.7°. A circle denotes the location detected from KOMPSAT EOC only.

3. 산사태 예측 모델 작성

1) 산사태 예측의 기본 가정 및 이론적 배경

산사태와 관련된 공간정보가 구축되게 되면, 구축된 정보를 유의미하게 통합하여 분석할 수 있는 통합 모델이 필요하다. GIS 기반의 공간통합 모델은 크게 목표유도형(Target-driven) 방법과 자료유도형(Data-driven)의 두가지로 나눌 수 있다 (Bonham-Carter, 1994). 목표유도형 방법은 전문가의 지식에 기반하여 통합에 사용되는 매개변수를 정하는 방법이며, 자료유도형 방법은 기존에 알려진 목적 사건에 대한 자료가 있을 때, 이를 이용하여 공간통합에 사용되는 각 주제도의 매개변수를 정하는 방법이다.

기본적으로 산사태 예측 모델을 작성하기 위해서는 다음과 같은 가정이 필요하다.

첫째, 산사태의 특징은 각 주제도에 의해 그 속성이 정의될 수 있다.

둘째, 미래의 산사태는 이미 일어난 산사태와 동일한 조건하에서 일어날 것이다.

위의 가정하에 A는 전체 연구 지역이고, L_k 를 k 번째 주제도라고 할 때, L_k 의 속성은 $\{1, 2, \dots, n_k\}$ 로 나타내어진다. 이 때, L_k 는 n_k 개의 속성값을 가지고 있다. 또한, 한 점 p에 대해 $V_k(p)$ 는 k번째 주제도

의 p점에서의 속성을 나타낸다. A의 한 픽셀 p에서 활동 유형의 미래의 산사태에 대해 영향을 받을 확률을 계산하기 위해 모든 p에서 m개의 속성이 있을 때, p가 영향을 받을 함수를 정의해야 한다.

본 연구에서는 기존 산사태 발생위치와 공간자료와의 상관성을 이용할 수 있는 자료유도형 통합 방법으로 결합 조건부 확률 모델(joint conditional probability model)과 certainty factor 모델을 이용하였다.

(1) 결합 조건부 확률 모델

“픽셀 p는 미래에 산사태가 일어날 것이다.”라는 명제 F_p 하에 A안에 있는 한 픽셀 p에 대하여 산사태에 영향을 받을 확률은 다음과 같이 정의할 수 있다.

$$\text{Prob}\{F_p | v_1(p), v_2(p), \dots, v_m(p)\} \quad (1)$$

또한, “ S_p : p는 과거 산사태에 의해 영향을 받았다.”라는 명제하에서 알려진 기존의 산사태에서의 결합 조건부 확률은

$$\begin{aligned} & \text{Prob}\{S_p | c_1, c_2, \dots, c_m\} \\ &= \text{Prob}\{p \in S_p \in \bigcap_{k=1}^m A_{k c_k}\} \quad (2) \\ &= \text{size of } \bigcap_{k=1}^m (S \cap A_{k c_k}) / \text{size of } \bigcap_{k=1}^m A_{k c_k} \end{aligned}$$

로 표시된다. 여기서 A_{km} 는 k 번째 주제도를 나타내고 m_k 는 k 번째 주제도 속성의 개수이다. 또한, 위의 식에서의 c_k 는 k 번째 주제도에서 p 의 속성을 나타낸다.

미래의 산사태 발생은 기존 산사태 발생과 주제도의 관계에 의해 유추될 수 있다는 가정과 조건부 독립 가정 및 베이시안 룰(Bayesian combination rule)에 의해 최종적으로 다음 식에 의해 산사태 발생 확률이 계산된다.

$$\begin{aligned} & \text{Prob}\{F_p|c_1, c_2, \dots, c_m\} \\ &= \frac{\text{Prob}\{c_1\} \dots \text{Prob}\{c_m\}}{\text{Prob}\{c_1, \dots, c_m\}} \text{Prob}\{S_p\} \\ & \frac{\text{Prob}\{S_p|c_1\}}{\text{Prob}\{S_p\}} \dots \frac{\text{Prob}\{S_p|c_m\}}{\text{Prob}\{S_p\}} \quad (3) \\ &= \frac{S}{\text{size of}(\bigcap_{k=1}^m A_{k c_k})} \frac{S_{1c_1}}{S} \dots \frac{S_{m c_m}}{S} \end{aligned}$$

(2) Certainty Factor 모델

Certainty factor(CF) 추정 방법은 전문가 시스템에서 주로 사용되며, 초기 확률에 대한 certainty의 측정방법이다. 이 방법은 예측 대상에 대한 영향요인이 많은 공간정보를 사용할 때 유효한 것으로 알려져 있다(Lee et al., 1999).

입력주제도 중 k 번째 주제도의 픽셀 p 에서의

certainty factor $CF_k(p)$ 는 픽셀 p 에서의 증거 $v_k(p)$ 가 주어지지 않은 경우로부터, 주어진 경우 까지, 픽셀 p 가 특정한 주제를 포함한다는 명제의 certainty의 변화로써 정의된다(Chung and Fabbri, 1993).

일반적으로 CF값은 -1에서 +1의 범위를 가지며, 양의 CF 값은 $v_k(p)$ 가 관찰된 이후에 명제의 certainty의 증가를 의미하며, 음의 CF값은 certainty의 감소를 의미한다.

CF는 확률의 함수로 주어지는데 식 (4)와 같다(Heckerman, 1986).

$$\begin{aligned} CF_k(p) &= \frac{\text{Prob}\{F_p|v_k(p)\} - \text{Prob}\{F_p\}}{\text{Prob}\{F_p|v_k(p)\}(1 - \text{Prob}\{F_p\})} \\ & \text{if } \text{Prob}\{F_p|v_k(p)\} > \text{Prob}\{F_p\} \\ &= \frac{\text{Prob}\{F_p|v_k(p)\} - \text{Prob}\{F_p\}}{\text{Prob}\{F_p\}(1 - \text{Prob}\{F_p\})} \quad (4) \\ & \text{if } \text{Prob}\{F_p|v_k(p)\} < \text{Prob}\{F_p\} \end{aligned}$$

위의 정의에 따르면, CF는 조건부확률이 초기 확률과 같으면 0이고, 조건부 확률이 초기확률과 차이가 있을 경우 절대값이 증가한다. 따라서 CF는 오로지 초기 확률에 관한 certainty의 측정이다.

2) 산사태 예측 모델 작성 결과

산사태 예측 모델은 Fig. 4의 절차에 의해 작성

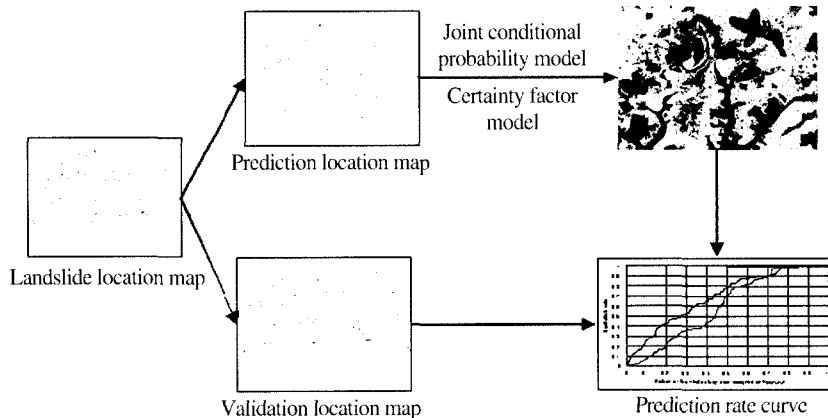


Fig. 4. Quantitative validation procedure in this study.

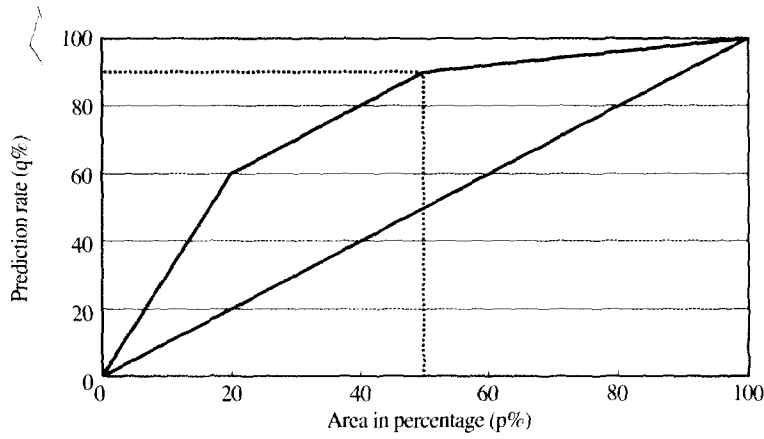


Fig. 5. Prediction rate curve.

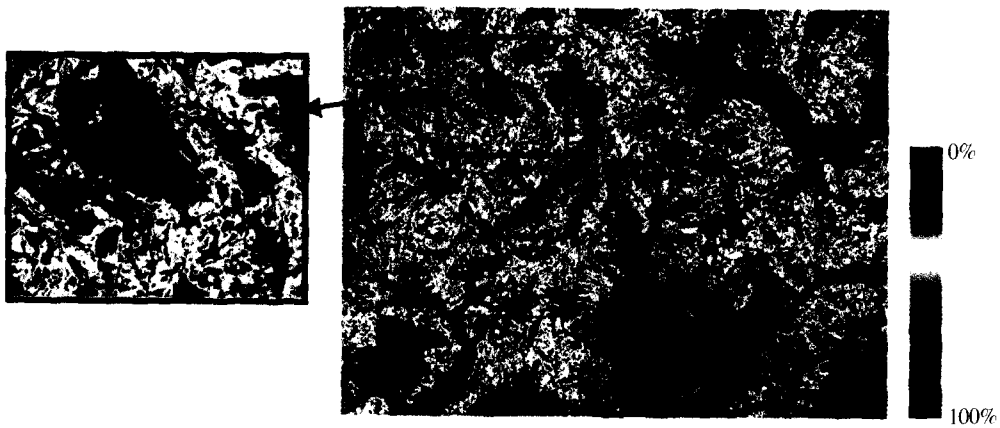


Fig. 6. Joint conditional probability prediction result. Black box is a detailed view of areas of which results are different from two models, and black polygons denote the validation sets (179 landslides) which are not used in construction of prediction model.

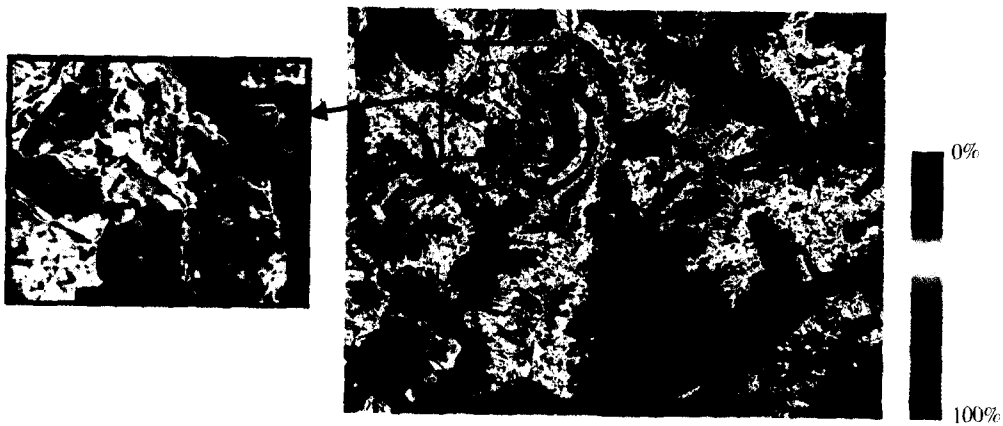


Fig. 7. Certainty factor prediction result. Black box is a detailed view of areas of which results are different from two models, and black polygons denote the validation sets (179 landslides) which are not used in construction of prediction model.

하였다. 359개의 산사태 발생위치를 무작위적으로 두 개의 그룹으로 나누어서 우선 180개는 통합에 사용하였고, 179개는 검증(validation)에 사용하였으며, 반대로 검증용 자료를 예측용 자료로, 그리고 예측용 자료를 검증용 자료로서 사용하였다.

통합 방법으로는 앞에서 설명한 결합 조건부 확률 모델과 certainty factor 모델을 사용하였다. 두 모델에서 초기 확률은 전체 연구지역에 대한 기존 산사태 발생의 면적 비율로 계산하였다. 결합 조건부 확률의 경우 값이 0에서 1, certainty factor의 경우 값이 -1에서 1까지 나오게 되는데, 두가지 방법으로부터 얻어진 결과를 비교하기 위하여 rank를 이용하여 결과를 도시하였다. 이를 위하여 예측 결과를 최대값으로부터 내림차순으로 정리한 후, 상대적 순위(rank)를 부여하였다. 최대값의 경우 1 혹은 0%, 최소값의 경우는 1/(전체 픽셀 개수) 혹은 100%로 나타나게 되며, %값이 작을수록 산사태 위험성이 높은 것을 의미한다.

위의 과정을 통해 얻어진 산사태 예측 결과를 보게 되면(Fig. 5, Fig. 6), 예측용 자료와 검증용 자료를 바꾸었을 경우 전반적인 양상에 변화는 없었으며, 통합에 사용된 두 방법에 의한 결과는 기본적으로 비슷한 양상을 보여준다. 결합 조건부 확률 모델의 경우 연구지역의 동쪽에 산사태 위험도가 높은 지역이 많이 나타나는 반면, certainty factor 모델의 경우 연구지역의 북쪽과 서쪽에 산사태 위험도가 높은 지역이 나타나고 있다. 통합에 사용되지 않은 검증 위치를 비교하게 되면, 특히 박스 부분의 경우 certainty factor 모델이 산사태 발생 지역에 위험도가 높은 지역이 집중적으로 나타나고 있다.

4. 예측비율곡선을 이용한 예측 모델 비교

1) 예측비율곡선

현재까지 연구의 대부분의 공간통합 모델의 개

발 및 적용에 주안점을 두어 진행되었으며, 예측 처리 결과의 정량화 검증에 대해서는 연구가 진행 중이다. 전통적인 검증 방법은 공간통합 모델을 작성하는데 사용한 산사태 발생자료를 가지고 예측 결과의 예측성을 비교하는 것이 대부분인데, 보다 엄밀한 의미의 검증을 위해서는 공간통합 모델을 작성하는데 사용하지 않은 산사태 발생자료를 가지고 예측 결과의 예측성을 비교해야 한다.

본 연구에서는 Chung and Fabbri (1999)에 의해 제안된 예측비율곡선(prediction rate curve)를 사용하였다. 예측비율곡선에서 x축값은 전체 연구지역의 전체면적에 대한 모델작성을 통해 계산된 특정 결과값까지의 면적의 비율을, y축값은 0%부터 특정 x값까지에 포함된 산사태 발생수의 전체에 대한 누적 비율을 의미한다. 예를 들어, Fig. 5의 경우 상대적인 산사태 발생 확률이 높은 상위 20%를 임의로 선택할 경우 그 부분에서 전체 산사태의 60%가 포함된다는 것을 나타내며, 50%를 임의로 선택할 경우 산사태의 90%가 포함된다는 것을 나타내고 있다. 따라서, 적은 면적에 많은 검증위치도의 산사태가 위치하여 예측이 잘 된 결과일수록 곡선은 기울기가 가파르고 곡선의 모양이 좌상으로 치우친 모양을 하게 된다.

2) 예측비율곡선 적용 결과

시각적인 비교로 certainty factor 모델의 예측 능력이 더 나은 것으로 나타났으나, 이를 정량적으로 표현하기 위하여 예측비율곡선을 작성하였다. 예측비율곡선을 작성함으로써, 본 연구지역의 상대적인 두 예측 모델의 예측능력 비교가 가능해진다.

예측비율곡선은 2가지의 경우로 나누어서 작성을 하였다. 우선 예측 결과의 특정 클래스에 포함되는 산사태의 픽셀수의 전체 산사태 픽셀수에 대한 비율로부터 계산하였다. 또다른 경우는 산사태 예측 결과의 특정 클래스에 포함되는 산사태의 비율을 다양하게 바꾸어 가면서 계산하였다. 일반적으로 산사태 예측 결과의 특정 클래스에 산사태가 걸쳐 있을 경우, 걸쳐 있는 비율은 한 개의 픽셀

부터 전체 픽셀까지 다양하다. 본 연구에서는 예측 결과의 특정 클래스에서 개별 산사태의 1개의 픽셀, 10% 및 20%가 각각 모두 포함이 되어야 예측이 되었다고 간주하여 예측비율을 계산하였다. 본 연구지역에서 발생한 산사태의 유형은 특별한 경우를 제외하고 고도가 높은 부분에서 시작되어 사면 아래로 유동된 흐름형(flow type)이어서 고도가 높은 부분으로부터 한 픽셀, 10%, 20%를 고려하였다. 이 경우 전체 픽셀수에 대한 비율로부터 계산이 되는 것이 아니라, 전체 산사태 발생 개수에 대한 비율로부터 계산이 수행된다.

359개의 산사태 발생위치를 무작위적으로 두 개의 그룹으로 나누었을 때, 우선 180개를 통합에 사용하고, 179개는 검증(validation)에 사용한 경우와, 반대로 검증용 자료를 예측용 자료로, 그리고 예측용 자료를 검증용 자료로서 사용한 경우는 Table 2과 Table 3에 제시하였다. 이 두 개의 경우를 평균하여 최종적인 예측비율곡선을 작성하였다(Fig. 8, Fig. 9).

픽셀수를 계산하여 예측비율곡선을 작성하였을 경우(Fig. 8), 조건부 확률 모델에서 상대적인 산사태 발생 확률이 높은 상위 10% 및 20%를 선택

Table 2. Four prediction rates of the prediction map w.r.t the 179 landslides.

rates	Proportion of 1444 pixels of landslides		20% of the scarps of 179 landslides		10% of the scarps of 179 landslides		1 pixel of the scarps of 179 landslides	
	conditional probability model	certainty factor model	conditional probability model	certainty factor model	conditional probability model	certainty factor model	conditional probability model	certainty factor model
0-5%	0.024	0.146	0.062	0.123	0.156	0.270	0.227	0.272
0-10%	0.07	0.262	0.130	0.178	0.334	0.395	0.362	0.405
0-15%	0.122	0.354	0.206	0.238	0.397	0.499	0.467	0.525
0-20%	0.177	0.427	0.281	0.271	0.528	0.564	0.574	0.589
0-25%	0.238	0.472	0.340	0.348	0.582	0.657	0.658	0.668
0-30%	0.319	0.524	0.421	0.458	0.668	0.686	0.708	0.781
0-35%	0.392	0.611	0.507	0.641	0.749	0.852	0.760	0.868
0-40%	0.487	0.648	0.601	0.785	0.781	0.892	0.805	0.915
0-45%	0.526	0.724	0.672	0.832	0.809	0.917	0.824	0.924
0-50%	0.651	0.792	0.734	0.868	0.833	0.926	0.842	0.936

Table 3. Four prediction rates of the prediction map w.r.t the 180 landslides.

rates	Proportion of 1640 pixels of landslides		20% of the scarps of 180 landslides		10% of the scarps of 180 landslides		1 pixel of the scarps of 180 landslides	
	conditional probability model	certainty factor model	conditional probability model	certainty factor model	conditional probability model	certainty factor model	conditional probability model	certainty factor model
0-5%	0.058	0.143	0.121	0.140	0.256	0.241	0.373	0.286
0-10%	0.134	0.263	0.199	0.236	0.377	0.399	0.48	0.411
0-15%	0.224	0.329	0.258	0.328	0.506	0.511	0.557	0.526
0-20%	0.278	0.392	0.343	0.354	0.576	0.573	0.63	0.598
0-25%	0.338	0.491	0.409	0.405	0.66	0.659	0.712	0.693
0-30%	0.403	0.576	0.514	0.493	0.696	0.701	0.739	0.715
0-35%	0.482	0.641	0.564	0.579	0.763	0.752	0.78	0.77
0-40%	0.531	0.679	0.651	0.687	0.79	0.785	0.814	0.796
0-45%	0.591	0.707	0.736	0.730	0.821	0.809	0.835	0.82
0-50%	0.717	0.724	0.776	0.779	0.835	0.824	0.85	0.837

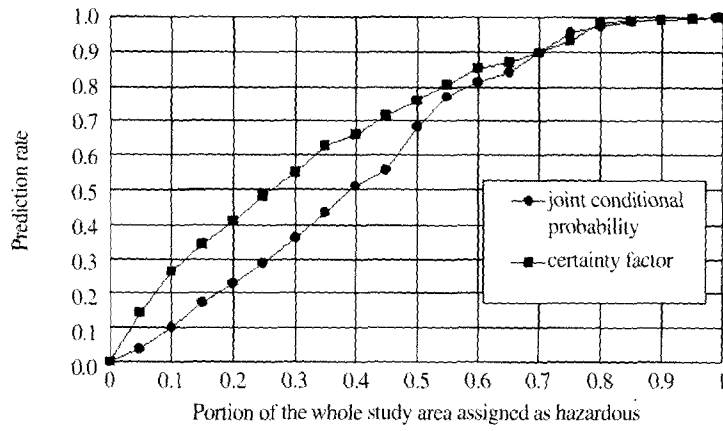
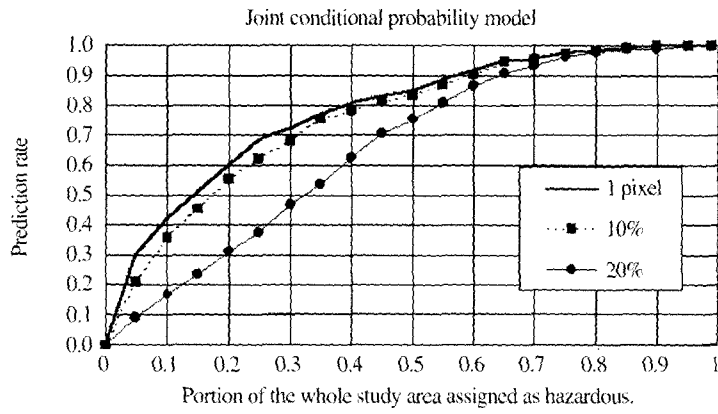
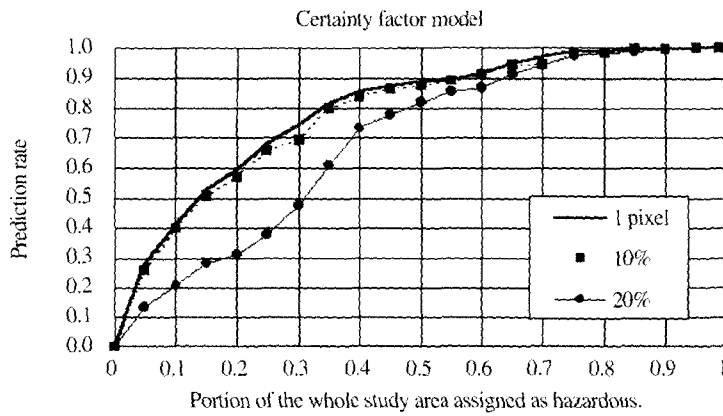


Fig. 8. Comparison of joint conditional probability prediction model and certainty factor prediction model w.r.t proportion of landslides pixels.



(a)



(b)

Fig. 9. Prediction rate curve for (a) joint conditional probability model, and (b) certainty factor model w.r.t various portion of the scarps.

하였을 경우 전체 산사태의 10% 및 20% 정도를 각각 설명할 수 있지만, certainty factor 모델의 경우 28% 및 40% 정도를 설명할 수 있는 것으로 나타났다. 산사태 발생 개수에 대한 비율로부터 계산된 결과의 경우, 결합 조건부 확률 모델(Fig. 9 (a))은 산사태 발생위치의 상위 10%가 포함되었을 때 예측이 되었다고 간주할 경우, 산사태 발생 확률이 높은 상위 20% 지역을 선정하였을 때 산사태 약 60%가 포함되었으며, certainty factor 모델(Fig. 9 (b))의 경우에도 비슷하게 나타났다.

5. 토의 및 결론

본 연구에서는 경기도 장흥지역을 대상으로 예측비율곡선을 이용한 산사태 예측 모델의 정량적 분석을 시도하였다. 우선 위성영상과 현지조사를 통해 산사태 위치를 파악하였으며, 산사태 발생과 관련이 있는 공간정보를 GIS 환경에서 구축하였다. 산사태 예측 모델을 작성하기 위하여 결합 조건부 확률 모델과 certainty factor 모델을 적용하였다. 기존 연구의 경우 산사태 발생 자료를 사용하여 예측 모델을 작성하고, 모델 작성에 사용된 산사태 발생 자료로 다시 추정 능력을 비교하였으나, 본 연구에서는 기존 산사태 발생위치를 공간적으로 구분하여 사용함으로써 예측 모델을 정량적으로 비교하였다.

본 연구에서 작성한 산사태 예측 모델과 같은 예측도 작성에서 한 시기에 동일 조건하에서 같은 원인으로 발생하는 사건을 추론할 때 어떠한 방법을 적용하던지 지역별로, 사건별로 정확한 추론은 불가능하다. 이러한 이유에서 현재의 산사태와 미래의 산사태는 동일조건하에서 발생한다는 가정하에 발생한 산사태의 일부를 발생하지 않은 것으로 가정하여 이에 대한 예측 가능성의 정량적 분석을 수행하였다. 본 연구지역에서의 결과는 certainty factor 모델이 결합 조건부 확률 모델보다 일반적으로 우수하다는 것은 아니며, 보다 많은

실험적 사례 연구를 수행할 필요가 있다.

GIS 기반 공간통합에서 고려해야하는 부분이 통합 모델의 선정과 오차 분석이다. 본 연구에서는 결합 조건부 확률 모델과 certainty factor 모델을 사용하였는데, 퍼지 연산 등의 다른 모델도 적용가능하다. 본 연구에서 사용한 예측비율곡선의 작성을 통하여 다른 모델과의 비교가 가능하며, 앞으로 여러 공간통합방법의 비교 연구 뿐만 아니라, 통합 목적, 자료의 종류에 따라 적절한 방법을 선택할 수 있는 근거를 마련하는 연구가 필요할 것으로 생각된다. 또한 자료 통합을 통한 예측도 작성에서는 통합방법론도 중요하지만, 실질적인 오차 분석도 중요한 부분이다. 보통 오차 분석은 임의의 오차를 포함하고 있는 실험자료를 통한 예비 연구가 필요한데 이에 대한 고려가 필요할 것으로 생각된다.

사 사

본 연구에 많은 조언을 해주신 캐나다 지질조사소의 정창조 박사님께 감사 드립니다. 그리고 논문의 내용을 보완하는데 도움을 주신 익명의 심사자분들께도 감사 드립니다.

참고문헌

- 김상규, 1988, 사면안정(I), 대한토목공학회지, 4(1): 90-101.
- 김윤중, 김원영, 유일현, 1992, 광역 지질재해분석(산사태)을 위한 GIS 활용, 대한지질공학회 지, 2(2): 131-140.
- 김원영, 이사로, 김경수, 채병근, 1998, 지형특성에 따른 산사태의 유형 및 취약성, 대한지질공학회지, 8(2): 115-130.
- 신진수, 2001, 다목적실용위성과 공간정보를 이용한 산사태 예측 기법 비교연구, 고려대학교 석사 학위 논문

- 이사로, 1999. GIS를 이용한 산사태 취약성 분석 기법 개발 및 적용 연구. 연세대학교 박사 학위 논문
- 홍원표, 1990. 사면안정(VIII). 대한토질공학회지, 6(3): 88-98.
- Bonham-Carter, G. F., 1994. *Geographic Information Systems for Geoscientists: Modeling with GIS*, Pergamon.
- Chung, F. C., and Fabbri, A. G., 1993. The representation of geoscience information for data integration, *Nonrenewable Resources*, 2(2): 122-139.
- Chung, C.F., and Fabbri, A.G., 1998. Three Bayesian prediction models for landslide hazard. In, A. Bucciantti, ed., *Proc. of International Association for Mathematical Geology Annual Meeting (IAMG98)*, Ischia, Italy, October 3-7, 1998, 204-211.
- Chung, F. C., and Fabbri, A. G., 1999. Probabilistic prediction models for landslide hazard mapping, *Photogrammetric Engineering and Remote Sensing*, 65(12): 1389-1399.
- Chung, C.F., and Fabbri, A. G., 2001. Validation of spatial prediction models for landslide hazard mapping, Submitted to *Natural Hazard* for publication.
- Heckerman, D., 1986. Probabilistic interpretations for MYCIN's certainty factors, in Kanal, L. N., and Lemmer, J. F., eds., *Uncertainty in Artificial Intelligence*, Elsevier, 167-196.
- Lee, K., N.-W. Park, K.-H. Chi, and B.-D. Kwon, 1999. GIS-based spatial integration and statistic analysis using multiple geoscience data sets : a case study for mineral potential mapping, *Korean Journal of Remote Sensing*, 15(2): 91-105.