

인공 신경망을 이용한 토석류 퇴적 모델 파라미터 추정

허경용[○], 박충식^{*}, 이창우^{**}, 윤호중^{**}

[○]동의대학교 전자공학과

^{*}영동대학교 스마트IT학부

^{**}국립산림과학원 산림방재연구과

e-mail: hgycap@deu.ac.kr[○], leciel@yd.ac.kr^{*}, {leecw, yount}@forest.go.kr^{**}

Parameter Estimation for Debris Flow Deposition Model Using Artificial Neural Networks

Gyeongyong Heo[○], Choong-Shik Park^{*}, Chang-Woo Lee^{**}, Ho-Joong Youn^{**}

[○]Dept. of Electronic Engineering, Dong-Eui University

^{*}Dept. of Smart Information Technology, Youngdong University

^{**}Dept. of Forest Conservation, Korea Forest Research Institute

● 요약 ●

토석류 퇴적 모델은 토석류에 의한 피해지 예측을 위해 그 효용성이 입증된 모델이지만 이를 이용하기 위해서는 몇 가지 파라미터를 필요로 한다. 파라미터를 자동으로 추정하기 위한 방법은 여러 가지가 있지만 토석류에 의한 피해지 예측을 위한 데이터는 충분히 양을 확보하기가 어려우므로 기존의 학습 기법을 적용하는데 어려움이 있다. 본 논문에서는 인공 신경망을 학습시키는 과정에서 기존 샘플로부터 의사 샘플을 생성하고 이를 학습에 사용함으로써 보다 안정적인 학습이 가능한 의사 샘플 신경망을 제안하였다. 제안한 의사 샘플 신경망은 해공간을 평탄화시킴으로써 잘못된 국부 최적해에 빠질 확률을 줄여주고 따라서 보다 안정적인 파라미터 추정이 가능하다는 사실을 실험을 통해 확인할 수 있다.

키워드: 토석류 퇴적 모델(debris flow deposition model), 파라미터 추정(parameter estimation), 신경망(neural network), 의사 샘플(pseudo sample)

I. 서론

우리나라의 토석류 피해는 대부분 집중 호우에 의해 발생되고 그 위치가 국지적으로 밀집된 양상을 보인다. 따라서 토석류 피해 위험지를 예측하고 이를 통해 사방구조물의 적지선정 및 경계 피난 체제 구축을 통해 피해의 규모를 줄일 수 있을 것으로 기대할 수 있다. 위험지 예측 및 구조물 위치 선정을 위해 이전 연구에서는 random walk model(RWM)을 이용하여 토석류 퇴적 모델을 구축하고 이를 통해 성공적인 결과를 얻을 수 있었다[1]. 하지만 RWM 모델에는 최적화시켜야할 세 개의 파라미터 값이 존재하며 이 값들은 경험적으로 결정되었다. 보다 넓은 지역에 토석류 모델을 적용하기 위해서는 파라미터 값을 자동으로 결정할 수 있는 방법이 필요하다.

파라미터 값을 추정하는 방법은 여러 가지가 있지만, 토석류 데이터의 경우 기존의 방법을 적용하기 위한 충분한 데이터를 확보하는데 어려움이 있어 적은 학습 샘플로 인해 신뢰성이 떨어지는 문제점이 있으며 이를 SSS (Small Sample Size) 문제라고 한다[2]. 이 논문에서는 인공 신경망을 사용함에 있어 SSS 문제를 완화하기

위해 의사 샘플을 생성하고 이를 학습에 사용하는 의사 샘플 신경망 (PSNN, Pseudo Sample Neural Network)을 제안한다 [3][4][5]. 제안한 PSNN은 의사 샘플을 통해 국부 최적해에 수렴하는 현상을 줄일 수 있으며 따라서 보다 안정적인 파라미터 추정이 가능해진다. 이러한 사실들은 실험 결과를 통해 확인할 수 있다.

II. 파라미터 추정

1. 요구 조건

토석류 데이터의 경우 측정 가능한 값과 파라미터 사이의 연관성이 명확하게 밝혀져 있지 않으며 이를 분석하기위한 충분한 데이터가 수집되어있지 않다. 현재 데이터는 계속해서 수집 중에 있으므로 향후 데이터는 지속적으로 늘어날 예정이다. 따라서 파라미터 추정 방법은 다음과 같은 요구 조건을 만족시켜야 한다.

- ① 입력과 출력의 관계를 명확하게 표현할 수 없는 경우 적용할 수 있어야 한다.

② 사용 가능한 데이터가 적은 경우에도 적용할 수 있어야 한다.
 ③ 추가되는 데이터를 기존 방법에 쉽게 반영할 수 있어야 한다.
 조건 ①과 ③은 신경망을 사용함으로써 만족시킬 수 있다[3].
 따라서 이 논문에서는 조건 ②를 해결하기 위해 의사 샘플 신경망
 을 제안한다.

2. 의사 샘플 신경망

의사 샘플 신경망(PSNN)에서 의사 샘플은 기존 샘플을 평균으
 로 가지는 가우스 분포를 이용하며 이 때 가우시안 분포는 중심
 극한 정리에 따라 선택되었다. N 개의 샘플 $X = \{x_1, x_2, \dots, x_N\}$
 이 주어진 경우 PSNN의 학습에 사용될 샘플은 식 (1)과 같이 나
 타낼 수 있다.

$$X_{PS} = \bigcup_{i=1}^N \left\{ x_j' \mid x_j' \sim N\left(x_i, \frac{\sigma^2}{N}\right), j=1, \dots, N_{PS} \right\} \quad (1)$$

이 때 σ^2 는 샘플 X 의 분산을 나타내고 N_{PS} 는 각 분포에
 서 생성되는 샘플의 개수를 나타낸다.

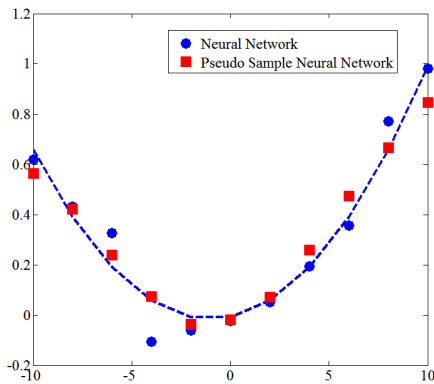


그림 1. 신경망과 의사 샘플 신경망의 학습
 Fig. 1. Training of a neural network and
 a pseudo sample neural network

PSNN을 이용하여 $y = \frac{1}{120}(x^2 + 2x - 1)$ 함수를 학습시킨
 결과가 그림 1로, 신경망에 비해 실제 함수를 보다 정확하게 추정
 하고 있음을 알 수 있다.

3. 토석류 데이터에 적용

토석류 데이터는 2개의 측정값으로부터 3개의 파라미터를 추정
 하여야 한다. 현재 확보된 데이터는 5개 지역에 대한 데이터뿐이
 므로 전체 데이터를 학습 및 테스트에 사용하는 방법을 사용하였
 다. 100회 반복 실험을 통한 파라미터 추정 결과를 표 1에서 확인
 할 수 있다.

표 1. 토석류 데이터에 대한 학습 결과
 Table 1. Training results using debris flow data

	신경망	의사 샘플 신경망
오류 평균	1.10	0.49
오류 분산	7.85	1.42

III. 결 론

이 논문에서는 학습 데이터의 크기가 작은 경우에도 학습의 안
 정성을 높일 수 있는 의사 샘플 신경망을 제안하였으며 그 유용성
 은 인위적인 데이터와 토석류 데이터를 이용한 실험에서 확인할
 수 있다. 토석류 데이터는 그 획득이 어려워 현재 적은 데이터만이
 확보된 상황이지만 지속적으로 데이터가 추가되고 있으므로 이후
 제안한 방법에 대한 유용성이 보다 면밀히 검증될 수 있을 것으로
 기대한다.

참고문헌

- [1] Chang-Woo Lee, Choongshik Woo, and Ho-Joong Youn, "Analysis of Debris Flow Hazard Zone by the Optimal Parameters Extraction of Random Walk Model – Case on Debris Flow Area of Bonghwa County in Gyeongbuk Province," Journal of Korean Forest Society Vol. 100, No. 4, pp. 664-671, 2011.
- [2] R.P.W. Duin, "Small sample size generalization," Proceedings of the 9th Scandinavian Conference on Image Analysis, pp. 957-964, 1995.
- [3] S. Haykin, "Neural Networks: A Comprehensive Foundation," 2nd ed. Prentice Hall, 1998.
- [4] T.G. Niel, T.R. McVicar, and B. Datt, "On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification," Remote Sensing of Environment, Vol. 98, No. 4, pp. 468-480, 2005.
- [5] C.M. Bishop, "Pattern Recognition and Machine Learning," 2nd ed. Springer, 2007.