

## Neuro-Fuzzy 초기화 함수 및 학습자료 구성에 따른 도시침수 위험기준 예측 모델 개선

### Improvement of Urban Flood Alert Criteria Prediction Model based on Neuro-Fuzzy Initial Function and Training Data

강호선\* · 조재웅\*\* · 이한승\*\*\* · 황정근\*\*\*\*

Kang, Hoseon\* · Cho, Jaewoong\*\* · Lee, Hanseung\*\*\* · Hwang, Jeonggeun\*\*\*\*

#### Abstract

In Korean metropolitan areas, the high density of residential and commercial buildings, highly impervious surfaces, and steep slopes contribute to floods that can occur within a short duration of heavy rainfall. To prepare for this, advance warning measures based on accurate flood alert criteria are needed. In our previous study, we demonstrated the applications of a Neuro-Fuzzy model that considers the characteristics of the basin to predict flood alert criteria in areas with no damage. However, as the number of learning materials are low, at 27, the evaluation and verification of the model has not been sufficiently accomplished, and its application is limited. Therefore, in this study, we propose an improved model based on the initializing function of the Neuro-Fuzzy algorithm, the construction of training data, and preprocessing. Compared to the existing model, the improved model reduced the average error by 48.1%~65.4% and the RMSE by 50.7%~60.1%. The new model, when applied to actual floods, showed an improvement of 0.7%~19.1% in accuracy.

**Key words** : Neuro-Fuzzy, Urban Flood, Training Data, Alert Criteria

#### 요 지

한국의 대도시는 주거 및 상가의 밀집도와 불투수율이 높고 유역경사가 급한 지형적 특징으로 집중호우 발생 시 단시간에 침수가 발생하여 대응시간이 부족하므로 사전에 대비하기 위해서는 정확한 침수위험기준을 활용한 침수 예경보가 필요하다. 선행연구에서는 피해이력이 없는 지역에 대한 침수위험기준을 예측하기 위해 유역특성을 고려한 Neuro-Fuzzy 모델을 개발하였으며, 실제 침수사상과 비교하여 적용성을 확인하였다. 그러나 학습자료의 수가 27개에 불과하여 모델 평가와 검증이 충분히 이루어지지 않았으며, 적용범위에도 한계가 있었다. 본 연구에서는 Neuro-Fuzzy 알고리즘의 초기화함수, 학습자료 추가 구축 및 전처리를 통해 모델을 개선하였다. 개선된 모델은 기존 모델에 비해 평균오차는 48.1~65.4%, RMSE는 50.7~60.1% 개선되었으며, 실제 침수사상과 비교한 결과에서도 약 0.7~19.1% 정확성이 개선되는 것으로 나타났다.

**핵심용어** : 뉴로-퍼지, 도시침수, 학습자료, 경보기준

\*정회원, 국립재난안전연구원 책임연구원(E-mail: hydro8487@korea.kr)

Member, Principal Researcher, National Disaster Management Research Institute

\*\*교신저자, 정회원, 국립재난안전연구원 시설연구사(Tel: +82-52-928-8175, Fax: +82-52-928-8199, E-mail: jwcho80@korea.kr)

Corresponding Author, Member, Research Officer, National Disaster Management Research Institute

\*\*\*정회원, 국립재난안전연구원 선임연구원

Member, Senior Researcher, National Disaster Management Research Institute

\*\*\*\*정회원, 국립재난안전연구원 연구원

Member, Researcher, National Disaster Management Research Institute

## 1. 서론

기후변화로 인해 지구의 평균기온이 지속적으로 증가하고 있으며, 이로 인한 이상기후 현상이 빈번하게 발생함에 따라 기후 관련 재해도 증가하고 있다. 특히, 홍수로 인한 재해가 가장 많이 발생하고 있으며, 피해규모도 가장 크게 나타났다. 2018년에도 자연재해 315건 중 홍수가 127건, 태풍 95건으로 자연재해의 약 70%를 차지하는 것으로 나타났다(CRED, 2019).

우리나라에서도 2008년~2017년까지 10년간 주요 자연재난 피해액을 살펴보면 태풍과 호우로 인한 피해액이 전체 피해액의 88.4%를 차지할 만큼 태풍과 호우는 많은 피해를 미치는 자연재난이다(MOIS, 2018). 주요피해사례로는 2010년 9월 21일 서울 및 인천지역 집중호우로 인한 침수, 2011년 서울 강남역, 사당역 침수와 우면산 산사태, 2012년 강남역 일대 침수, 2014년 부산 온천천 일대 침수, 2016년 태풍 “차바(Chaba)” 내습으로 인한 울산지역 침수, 2017년 장마기간 집중호우로 인한 인천 및 부산 침수피해 등이 있다. 이와 같이 대도시권에서 침수로 인한 피해가 큰 원인으로서는 우리나라의 대도시는 주거 및 상가의 밀집도와 불투수율이 높고 유역경사가 급한 지형·환경적인 특징으로 인해 집중호우 발생 시 매우 짧은 시간에 침수가 발생하여 대응할 시간이 부족하며, 차량 및 인구가 집중되어 있어 그 피해가 크게 발생하기 때문이다.

도시집에서 극한강우로 인한 침수피해 저감대책으로 배수펌프장 설치, 우수관 배수용량 확보 등의 구조적인 대책으로는 한계가 있으며, 사전예보 및 대피 등 비구조적인 대책이 함께 이루어져야 한다. 그러나 우리나라의 도시침수 저감대책은 구조적 개선 방법을 중심으로 이루어지고 있으며, 비구조적인 대책으로는 하천에서 수위계측자료를 활용한 홍수에·정보만 이루어지고 있다.

‘지역별 주요 재난대응 시나리오 및 기준 개발’ 연구를 통해 서울 강남구, 서초구, 관악구, 양천구, 강서구에 대한 피해이력기반의 침수위험기준 산정식을 개발하여 도시침수에·정보를 위한 기준을 제시하였으며, ‘지역별 도시침수 대응체계 기반 구축’연구에서는 공간적 단위를 세분화 하여 행정동, 배수분구 단위의 침수위험기준을 제시하였다. 또한 피해이력이 없는 지역에 침수위험기준을 예측하기 위한 Neuro-Fuzzy 모델을 활용한 유역특성기반의 침수위험기준 추정방법을 제시하고 추정 결과의 시범 적용을 통해 적용성을 확인하였다(NDMI, 2014, 2015, 2016, 2017).

Cho et al. (2018) 연구에서는 초기 모델 설계과정으로 초기화 함수를 개발자의 판단이 알고리즘의 구조 결정에 영향을 미치는 Grid Partition으로 하였으며, 서울 25개, 울산 2개 읍면동, 총 27개의 학습자료를 사용하여 모델을 설계하였다. 학습자료는 종속변수인 피해이력기반 침수위험기준 산정 결과와 독립변수인 도시지역의 유출에 영향을 주는

유역특성자료(유역경사, 불투수율, 관거밀도, 배수펌프장 수방능력, 빗물받이 밀도)를 사용하여 침수위험기준 예측 모델을 개발하였다. 그러나 학습자료로 서울과 울산지역의 유역특성만 사용하여 적용범위가 작고, 적용범위를 벗어난 경우 침수위험기준 예측이 불가능한 한계점도 확인하였다. 또한 학습자료의 부족으로 교차검증을 통해 모델을 검증하고 실제 침수가 발생한 지역의 침수사상과 예측된 침수위험기준을 비교하여 모델의 적용성을 확인하였으나, 모델에 대한 충분한 평가와 검증이 이루어지지 않았다.

본 연구에서는 제시된 유역특성기반의 침수위험기준 예측 모델의 초기화함수, 학습자료의 추가 구축 및 전처리 과정을 통해 모델 예측 결과의 정확성을 향상시키는데 목적이 있다.

## 2. 연구내용 및 방법

유역특성기반의 침수위험기준 예측 모델에 사용된 Neuro-Fuzzy는 인간의 추론능력을 근사한 퍼지이론과 학습 및 적응 능력이 있는 신경망이 결합되어 복잡성과 불확실성을 나타내는 비선형 시스템에서 효과적인 알고리즘이다. Neuro-Fuzzy 모델의 기본 구조도는 Fig. 1에 나타내었다. 입력변수로 불투수율, 유역경사, 관거밀도, 빗물받이 밀도를 사용하였으며, 결과변수로 피해이력기반 침수위험기준 추정 값을 사용하였다.

선행연구에서 제시된 유역특성기반의 침수위험기준 예측 모델 개선을 위해 Fig. 2와 같이 ① 초기화 함수와 ② 학습자료의 수, ③ 전처리 기법 등 모형의 구조와 학습자료의 구성을 다양하게 적용하여 비교하였으며, 모델 별 성능평가를 통해 최적 모델을 제시하였다. 또한 기존 모델과 개선된 모델을 실제 침수발생강우량과 비교하여 개선효과를 분석하였다.

먼저 Neruo-Fuzzy 알고리즘에서 퍼지구조를 결정하는 요소인 초기화 함수 구성방법에 따른 모델 정확성을 비교하기 위해 선행연구에서 적용된 Grid Partition과 새로운 기법인 Subtractive Clustering 방법을 적용하여 비교하였다.

여기서 Grid Partition과 Subtractive Clustering은 퍼지추론시스템(Fuzzy Inference System)의 초기화 방법으로 독립변수(유역경사, 불투수율, 관거밀도, 배수펌프장 수방능력, 빗물받이 밀도)와 종속변수(침수위험기준) 사이의 관계를 정의하는 초기 멤버십 함수를 설정하는 방법이다. Grid Partition은 각 독립변수와 종속변수의 전 범위를 일정한 간격의 멤버십 함수로 구성하고, 입·출력변수(독립변수-종속변수)를 Fig. 3(a)와 같은 격자 형태로 구성하여 각 구간에 연결된 독립변수와 종속변수의 관계를 멤버십 함수로 정의하는 방법이다. 반면에 Subtractive Clustering 기법의 경우 Fig. 3(b)와 같이 각 인자들이 밀집되어 있는 범위만을 대상으로 조건·결과 관계를 구성하는 방법을 선택하게 된다. 퍼지를

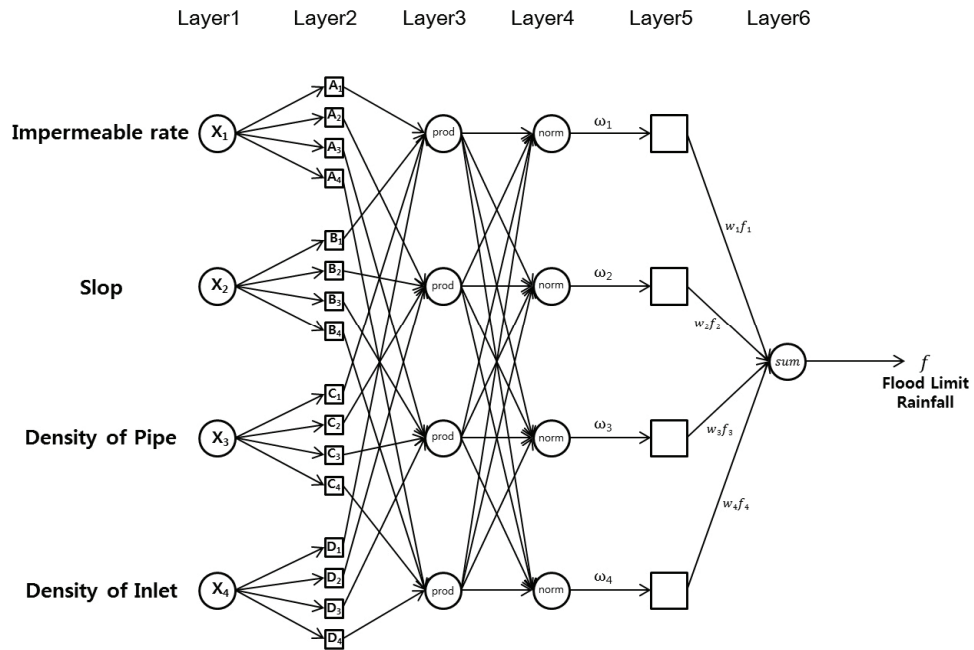


Fig. 1. Structure of Neuro-Fuzzy Model (Cho et al., 2018)

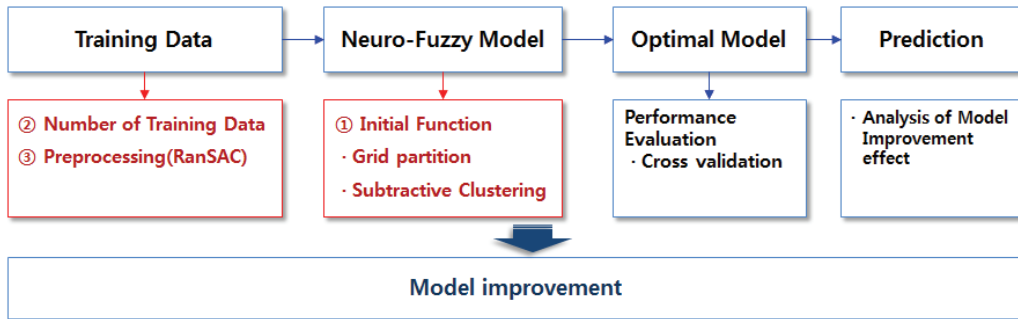


Fig. 2. Research Process

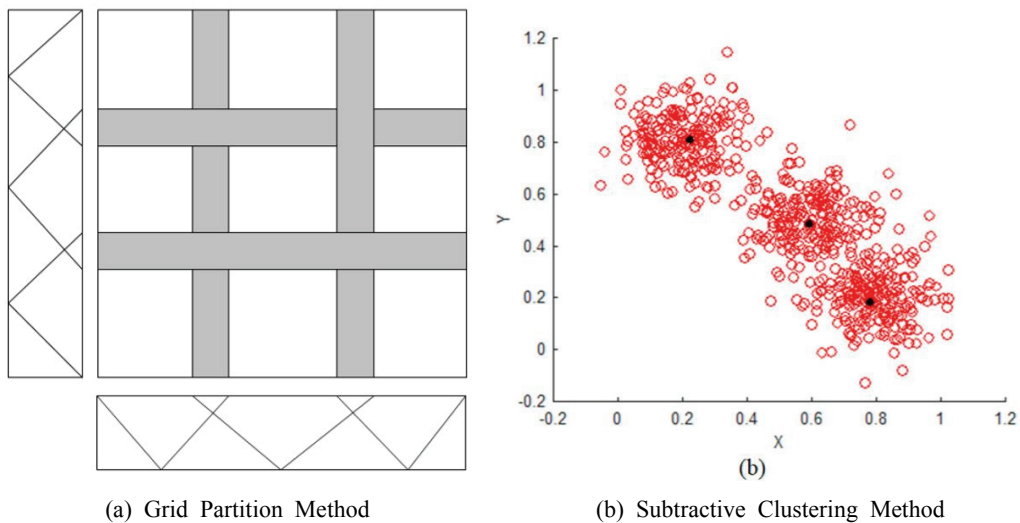


Fig. 3. Fuzzy Inference Process (Yeom and Kwak, 2018)

(조건-결과 관계식)의 수와 결과의 정확성과는 상관성이 없으며, Grid Partition을 통해 입·출력변수의 전 구간에 대한 퍼지관계식을 만들기보다는 Subtractive Clustering 기법을 사용하여 모형의 효율성 및 블랙박스모형의 장점을 최대한 활용하는 방법이 많이 사용되고 있다.

두 번째로 161개 학습자료를 추가로 구축하고 학습자료의 수를 다양하게 적용하여 모델 정확성을 비교하였으며, 마지막으로 정상범위를 벗어난 학습자료를 제거하기 위해 RanSAC 기법을 적용하여 전처리 과정을 거친 학습자료를 사용한 모델과 비교하였다.

RanSAC 기법은 거짓정보를 포함하는 일련의 데이터 집합으로부터 수학적 모델 인자들을 반복적인 작업으로 예측하여 학습자료의 이상치(Outlier) 즉, 정상분포(Inlier)에서 벗어난 값을 학습자료에서 제외시키는 기법이다. 예를 들어 Fig. 4(a)와 같이 노이즈(Noise) 데이터를 포함한 학습자료 전체를 활용할 경우 정상범위의 데이터에서도 오차가 발생할 수 있다. 그러나 Fig. 4(b)와 같이 RanSAC 알고리즘을 활용하여 Outlier를 제거한 경우 정상범위의 데이터에서

정확한 결과를 얻을 수 있다.

마지막으로 초기화 함수 및 학습자료의 구성에 따른 최적 모델을 제시하고, 미세측 지역에 대한 침수위험기준 예측 결과를 실제 침수사상과 비교하여 모델의 적용성을 확인하였다.

도시침수 위험기준 예측 모델 개선 효과를 비교·분석하기 위해 Table 1과 같이 초기화 함수, 학습자료 수, RanSAC 기법 적용 여부에 따라 9개의 Case로 학습자료를 구성하였다.

모델의 평가방법으로는 학습데이터의 수가 작은 경우에도 검증이 가능한 교차검증 방법을 적용하였다.

Fig 5와 같이 전체 자료 수 N개 중 2개 자료를 검증(Test)자료로 사용하고, 나머지 N-2개의 자료를 학습(Train)에 사용하는 방법으로 N-1개의 데이터 셋으로 구성하였다. 예를 들면 Model A의 경우 27개 자료 중 25개를 학습자료, 학습에 사용되지 않은 2개를 검증자료로 사용하여, 총 26개의 데이터 셋으로 구성하였다.

각 모델간의 비교를 위해 평균오차와 Root Mean Square Error (RMSE)를 비교하여 모델을 평가하였다.

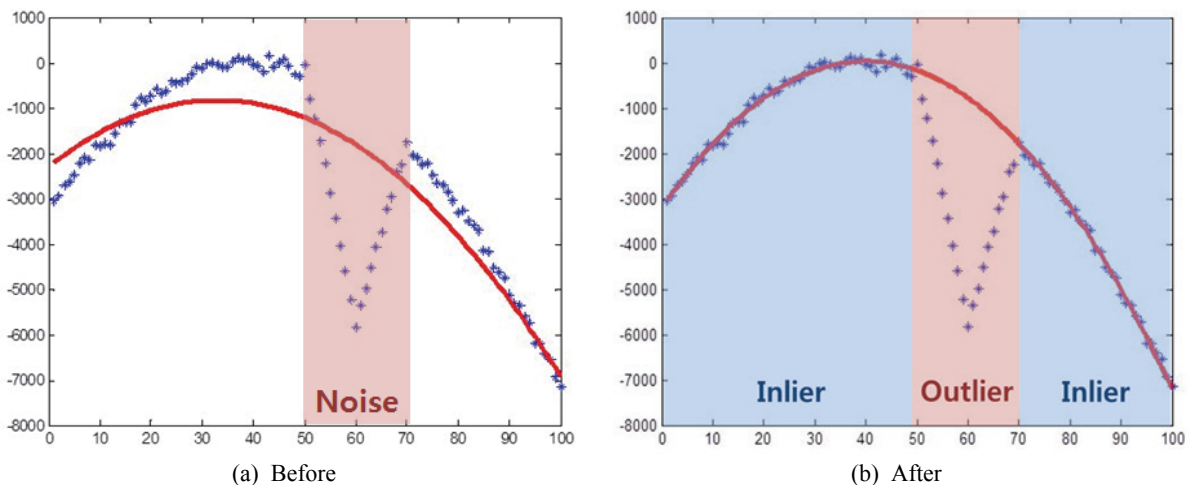


Fig. 4. Comparison Between Before and After RanSAC Algorithm

Table 1. Neuro-Fuzzy Model Training Data Composition

Case	Clustering Technique	Number of Training Data	Preprocessing
Model A	Grid Partition	27	-
Model B	Grid Partition	40	-
Model C	Grid Partition	119	-
Model D	Grid Partition	188	-
Model E	Subtractive Clustering	27	-
Model F	Subtractive Clustering	40	-
Model G	Subtractive Clustering	119	-
Model H	Subtractive Clustering	188	-
Model I	Subtractive Clustering	188	RanSAC

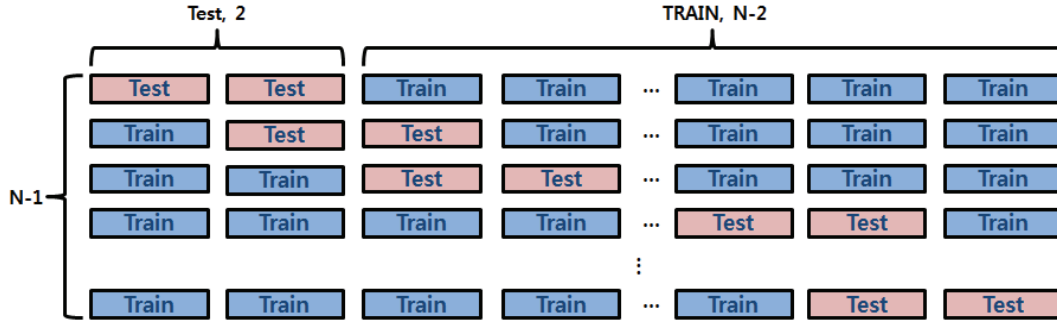


Fig. 5. Cross Validation

여기서, 평균오차는 Eq. (1)로 계산되며, Root Mean Square Error (RMSE)는 평균제곱근오차를 의미하며, 표준편차를 일반화시킨 값으로 실제값과 추정값의 차이를 나타내는 척도로 사용된다. RMSE는 Eq. (2)를 이용하여 산정하였다.

$$Error = \frac{1}{n} \sum_{i=1}^n (\hat{y} - y) \quad (1)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y)^2} \quad (2)$$

Eqs. (1), (2)에서  $\hat{y}$ 는 실제값을 의미하며,  $y$ 는 모델에 의해 예측된 추정값을 의미한다.

### 3. 초기화 함수 개선

초기화 함수는 Neuro-Fuzzy 알고리즘에서 Fuzzy 구조를 결정하는 함수로 조건에 따라서 모델의 예측력, 학습시간 등에 영향을 줄 수 있다. 선행연구에서는 침수위험기준 예측을 위한 Neuro-Fuzzy 알고리즘의 적용 가능성을 시험해보는 초기단계연구로 Fuzzy 멤버십 함수(Membership Function)의 형태 및 개수 등을 개발자가 직접 지정하고 모형의 학습에 사용하는 초기 멤버십 함수의 지정에 군집화 기법 중 Grid Partition 방법을 사용하는 등 개발자의 판단이 알고리즘의 구조 결정에 영향을 미치는 학습방법을 사용한 모델 연구를 수행하였다. 즉, 침수위험기준 예측에서 가장 중요한 피해

이력기반의 침수위험기준 추정 결과가 부족한 상황에서 모형의 정확성 및 적용성 향상 방안에 중점을 둔 연구를 수행하였다.

선행 연구에서 사용된 멤버십 함수는 Table 2와 같다.

본 연구에서는 Subtractive Clustering 기법을 적용하여 침수 위험기준 예측 모델을 재구성하였으며, Grid Partition 기법을 적용한 모델과 비교하여 정확성이 높은 기법을 모델에 적용할 것이다. Subtractive Clustering 기법의 경우 Influence, Quash Factor, Accept Ratio, Reject Ratio 4개의 파라미터가 있으며, Table 3과 같이 기본으로 제공하는 값을 사용하였다. 퍼지추론기법 학습과정에서는 기존 연구와 동일하게 Least-Square 기법과 Back-Propagation 기법을 병행하여 사용하는 Hybrid 기법을 사용하고, 10회 반복하여 학습하도록 하였다.

초기화 함수에 따른 개선 효과를 분석하기 위해 학습자료 수를 동일하게 하고 초기화 함수를 Grid Partition과 Subtractive Clustering 적용한 Table 1의 Model A, B, C, D와 Model E, F, G, H을 비교하여 Table 4와 Fig. 6에 나타내었다.

학습자료의 수가 작을 경우 Grid Partition이 Subtractive Clustering 보다 정확성이 높았으며, 학습자료가 증가한 경우 Subtractive Clustering 방법이 정확성이 높아지는 것으로 나타났다. 특히 학습자료 119개를 사용한 조건에서 Grid Partition을 사용한 모델의 평균오차가 크게 발생하였는데, 이는 종속변수인 침수위험기준은 동일하지만 독립변수가 되는 유역특성자료의 수치가 다른 학습자료 즉, 모델 학습에 부정적인 영향을 주는 학습자료가 다수 포함되어 있기 때문

Table 2. Grid Partition Parameter

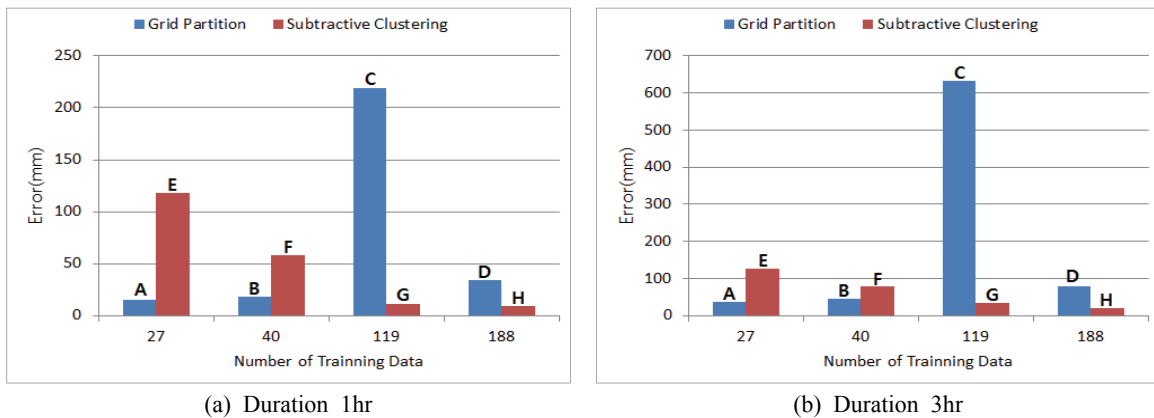
Initial Function	Membership Function	Number of Membership Function
Grid Partition	2-Variable Gaussian	5

Table 3. Subtractive Clustering Parameter Default Value

Influence Range	Quash Factor	Accept Ratio	Reject Ratio
0.5	1.25	0.5	0.15

**Table 4.** Model Improvement Effect by Initial Function

Division	Duration	Error by Initial Function (mm)	
		Grid Partition	Subtractive Clustering
Model A-E (27)	1hr	15.8	117.7
	3hr	36.1	125.7
Model B-F (40)	1hr	18.0	58.4
	3hr	45.0	77.3
Model C-G (119)	1hr	218.8	11.0
	3hr	631.9	35.0
Model D-H (188)	1hr	34.2	9.7
	3hr	79.8	19.8



**Fig. 6.** Error (mm) Comparison by Initial Function

인 것으로 판단된다. 반면에 Subtractive Clustering 방법은 이와 같은 문제를 잘 보정하는 것으로 나타났다.

#### 4. 학습자료 추가 구축

학습자료의 추가 구축을 위해 최근 5년('13~'17)간 침수 피해가 많은 지역 123개 시군구 1,784개 읍면동을 분석하여, 서울 8개, 대전 3개, 광주 1개, 인천 25개, 부산 47개, 경기도 105개, 강원도 1개, 경상북도 11개, 충청남도 4개 등 205개 읍면동에 대한 침수위험기준을 추정하였다. 침수위험기준 추정 결과는 Table 5와 같이 요약하였다.

침수위험기준 예측 모델 개발을 위한 학습자료 구축을 위해 각 지자체에 요청하여 최신의 지형공간정보자료를 수집하였으며, 관거밀도, 불투수율, 유역경사, 빗물받이 밀도 등 4개의 유역특성자료를 GIS분석을 통해 구축하였다. 피해이력기반으로 침수위험기준을 추정한 205개 읍면동 중 유역특성자료가 구축된 지역은 161개로 선행연구에서 사용된 학습자료 27개를 포함하여 총 188개를 학습자료를 구축하였다.

선행연구에서 주요매개변수로 사용했던 빗물펌프장의

수장능력은 객관적인 수치를 확보할 수 없을 뿐만 아니라, 빗물펌프장 시설의 유무에 따라 침수위험기준 예측결과에 영향이 크게 나타나므로 학습자료에 부적합 하다고 판단하여 유역특성 자료에서 제외하였다.

학습자료의 유역특성범위는 Table 6에 나타내었다. 새로운 지역에 대한 침수위험기준을 예측할 경우 예측하고자 하는 지역의 유역특성자료가 모델 학습자료에 사용된 범위를 벗어날 경우 예측결과가 발산하거나 큰 오차가 발생하는 등 신뢰할 수 없는 결과값이 주어지기 때문에 유역특성자료의 범위는 매우 중요하며, 모형을 통한 침수위험기준 예측시 사용된 학습자료의 적용 범위를 반드시 확인할 필요가 있다. 관거밀도는 0~0.5222, 불투수율은 1.2~100.0, 유역경사, 0.01~39.85%, 빗물받이 밀도는 0~0.0055의 범위를 갖는 것으로 나타났으며, 관거밀도와 빗물받이 밀도의 경우 하수시설이 구축되지 않거나 유역면적대비 매우 작은 비중을 차지하는 지역이 있어 최소값이 0인 유역이 발생하였다.

학습자료 수에 따른 모델 정확성 비교를 위해 초기화 함수를 Subtractive Clustering으로 하고 학습자료의 수를 각각 27개, 40개, 119개, 188개로 증가시키면서 학습한 Model E, F, G, H의 평균오차와 RMSE를 Table 7에 나타내었다.

**Table 5.** Result of Urban Flooding Limit Rainfall by Damage Records

City	District	Flooding Limit Rainfall by Duration (mm)	
		1hr	3hr
Seoul	Sillim-dong	71.8	126.1
	Gonghang-dong, Mok-3dong	82.0	175.4
Busna	Gadeokdo-dong, Daejeo-2dong, Noksan-dong	52.3	83.7
	Myeongji-dong	47.1	83.2
	Goejeong-1dong	43.3	73.9
Suwon	Godeung-dong, Gyo-dong and so on	70.8	129.8
	Guun-dong, Geumgok-dong and so on	59.5	96.0
Changwon	Gapo-dong, Gusan-myeon and so on	86.5	209.6
⋮	⋮	⋮	⋮

**Table 6.** Range of Basin Characteristics Parameter

Division	Density of Pipe	Impermeable Rate (%)	Slope (%)	Density of Inlet
Minimum	0.0000	1.20	0.01	0.0000
Maximum	0.0522	100.00	39.85	0.0055
Mean	0.0109	57.76	10.68	0.0007

**Table 7.** Model Improvement Effect by the Number of Training Data

Division		Duration	
		1hr	3hr
Model E (Subtractive Clustering, 27)	ERROR (mm)	117.7	125.7
	RMSE (mm)	345.0	191.5
Model F (Subtractive Clustering, 40)	ERROR (mm)	58.4	77.3
	RMSE (mm)	264.3	143.6
Model G (Subtractive Clustering, 119)	ERROR (mm)	11.0	35.0
	RMSE (mm)	13.7	44.7
Model H (Subtractive Clustering, 188)	ERROR (mm)	9.7	19.8
	RMSE (mm)	12.8	29.9

학습자료가 27개에서 40개로 증가할 경우 지속시간 1시간의 경우 평균오차는 50.3% 감소하였으며, 119개일 경우 90.7%, 188개일 경우 91.8% 감소하는 것으로 나타났다. 지속시간 3시간에서는 각각 38.5%, 72.2%, 84.2% 감소하였다. RMSE도 지속시간 1시간에서 23.4%~96.3%까지 감소하였으며, 3시간에서도 25.0%~84.4% 까지 감소하였다.

Figs. 7~8의 학습자료의 수 증가에 따른 평균오차와 RMSE를 지속시간별로 비교한 그래프를 통해 모델 학습자료의 수가 증가하면 평균오차와 RMSE는 감소하는 것을 확인할 수 있다. 따라서 학습자료 수가 증가함에 따라 모델의 정확성은 증가하며, 학습자료의 추가 구축을 통해 모델의 정확도를 향상시킬 수 있을 것으로 판단된다.

## 5. 데이터 전처리 기법 적용

선행연구에서는 학습자료가 27개에 불과하여 데이터 전처리가 불가능 하였다. 그러나 본 연구에서는 학습자료의 추가구축을 통해 총 188개의 학습자료를 확보하였고, 예측 결과에 부정적인 영향을 미치는 학습자료를 선별하기 위해 RanSAC기법을 적용하였다. 수행과정은 학습자료에서 임의의 개수를 정상분포(Inlier)로 가정하여 회귀모델을 구하고, 나머지 데이터들을 회귀모델과 비교하여 허용오차 내에 있는 학습자료를 Inlier로 포함하고 재구성된 Inlier를 이용하여 다시 회귀모델을 구하고, 회귀모델과 Inlier의 오차를 측정하여 허용오차 범위 내에 도달할 때까지 반복하였다. 그 결과 총 188개 학습자료 중 지속시간 1시간에서는 정상분

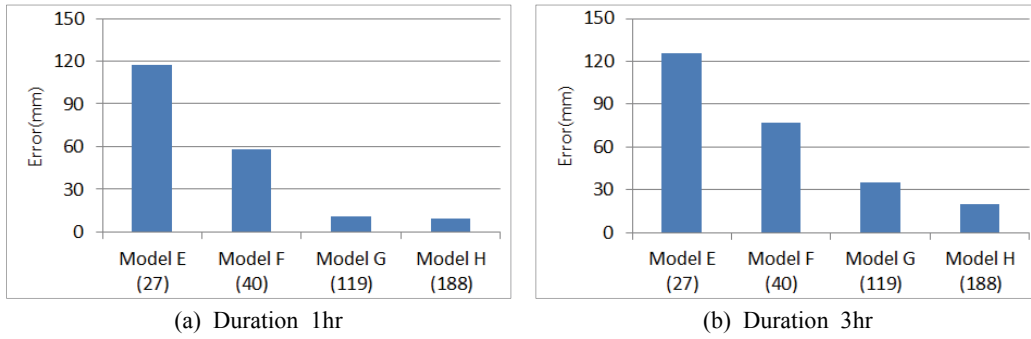


Fig. 7. Error (mm) Comparison by the Number of Training Data

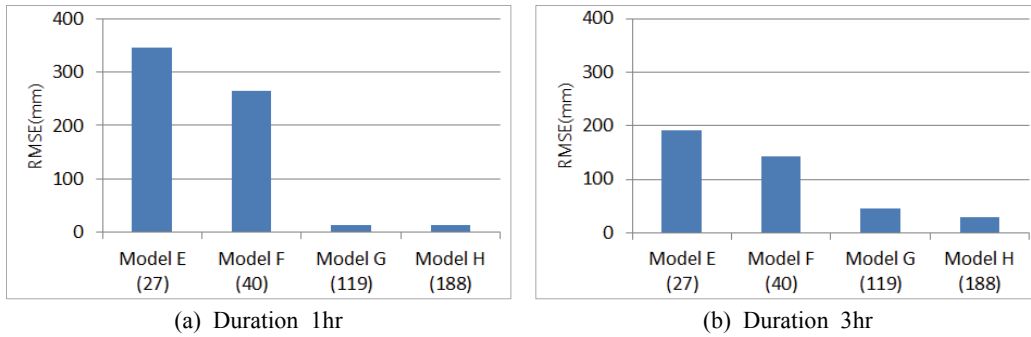


Fig. 8. RMSE (mm) Comparison by the Number of Training Data

포 범위를 벗어난 45개를 제외한 145개, 3시간에서는 90개를 제외한 98개를 학습자료로 사용하여 침수위험기준 예측 모델을 설계하였다.

RanSAC 기법을 적용한 Model I와 188개 학습자료를 모두 사용한 Model H를 비교하여 Table 8 및 Fig. 9에 나타내었다. 평균오차는 지속시간 1시간에서 1.5 mm 감소, 3시간에서는 7.3 mm 감소하는 것으로 나타났으며, RMSE도 각각 2.5, 11.7 감소하여 모델 정확도가 증가하는 것으로 분석되었다.

## 6. 모델 개선 효과 비교 및 적용성 검증

총 123개 시군구 1,784개 읍면동 중 피해이력기반의 침수 위험기준 추정이 불가능한 1,579개 읍면동의 유역특성자료를 수집·분석하고, Model I에 적용하여 침수위험기준을 예측하였다. 예측된 침수위험기준 결과 중 일부를 Table 9에 제시하였다.

본 연구에서 제시된 개선된 침수위험기준 예측 모델의 개선효과 및 적용성을 검토하기 위해 2017년과 2018년에 침수피해가 발생한 지역의 CCTV 영상자료를 수집하였다. 2017년에는 부산 강서구 녹산동에서 침수가 발생하였으며, 2018년에는 대전 유성구 전민동과 서구 둔산동, 울산 북구 효문동에서 침수가 발생하였다. 침수발생지역의 CCTV 영상을 수집·분석하여 침수발생 시간을 확인하였으며, 분석 결과를 Table 10에 나타내었다.

해당지역의 AWS 강우관측 자료를 통해 침수발생시간 당시 지속시간별 누적강우량을 분석하여 실제 침수발생강우량을 산정하였다. 실제 침수발생강우량을 침수위험기준의 참값으로 기존 모델로 예측된 침수위험기준과 개선된 모델로 예측된 침수위험기준을 비교하여 개선 효과 및 적용성을 검토하였다.

예측된 침수위험기준의 정확성을 검토하기 위해서는 실제 침수를 발생시킨 강우의 지속시간을 확인하여 동일한

Table 8. Result of Urban Flooding Limit Rainfall by Model

Model	Division	Duration	
		1hr	3hr
Model H (Subtractive Clustering, 188)	RMSE(mm)	12.8	29.9
	ERROR(mm)	9.7	19.8
Model I (Subtractive Clustering, RanSAC)	RMSE(mm)	10.3	18.2
	ERROR(mm)	8.2	12.5



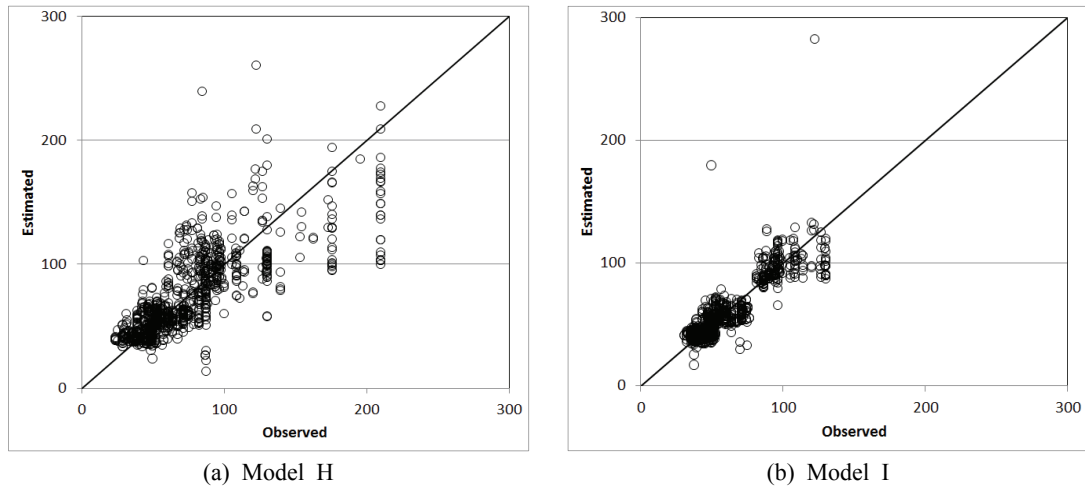


Fig. 9. Comparison of Model H - I

Table 9. Result of Urban Flooding Limit Rainfall by Model

City	District	Flooding Limit Rainfall by Duration (mm)	
		1hr	3hr
Busan	Garak-dong	63.3	115.1
	Gangdong-dong	63.2	121.9
	Dadae-2dong	45.1	98.5
Suwon-si	Gokseon-dong	65.2	113.0
	Gwonseon-2dong	63.7	107.8
Gwangju	Dosan-dong	61.9	105.4
	Donggok-dong	65.2	118.3
Changwon-si	Yangdeok-2dong	66.0	127.3
	I-dong	60.9	117.8
⋮	⋮	⋮	⋮

Table 10. Actual Flooding Time

Year	City	District	Flooding Date	Flooding Time
2017	Gangseo-gu, Busan	Noksan-dong	Sept. 11.	06:45
2018	Yuseong-gu, Daejeon	Jeonmin-dong	Aug. 28.	06:08
	Seo-gu, Daejeon	Dunsan-dong	Aug. 28.	06:14
	Buk-gu, Ulsan	Hyomun-dong	Aug. 26.	11:38

지속시간에 대해서 비교해야 한다. 따라서 침수를 발생강우의 지속시간을 확인하기 위해 지속시간 1시간, 3시간 강우에 대한 빈도분석을 하였으며, 빈도가 높은 강우사상에 의해 침수가 발생한 것으로 판단하였다. 실제침수발생강우량의 지속시간별 빈도분석결과를 Table 11에 나타내었다.

부산 녹산동은 지속시간 1시간에 2.4년 빈도, 3시간에 1.8년 빈도로 43.0 mm/hr 강우에 의해 침수가 발생하였으며, 대전 전민동과 둔산동은 1시간에 5.5년, 4.2년 빈도, 3시간에 2.9년, 3.2년 빈도로 각각 59.6 mm/hr, 55.9 mm/hr 강우에 의해 침수가 발생하였다. 울산 효문동은 지속시간 1시간에

3.0년 빈도, 3시간에 7.3년 빈도로 96.0 mm/3hr 강우에 의해 침수가 발생한 것으로 판단하였다.

부산 녹산동과 대전 전민동, 둔산동은 지속시간 1시간 강우, 울산 효문동은 지속시간 3시간 강우에 의해 침수가 발생한 것으로 확인되었으며, 기존 모델 예측결과와 개선된 모델 예측결과를 비교하여 Table 12에 나타내었다. 부산 녹산동은 43.0 mm/hr 강우에 의해 침수가 발생하였으며, 기존 모델은 52.3 mm/hr, 개선 모델은 44.1 mm/hr의 침수위험기준을 예측하여, 개선 모델이 19.1% 개선되는 것으로 나타났다.

**Table 11.** Frequency of Actual Flooding Rainfall by Duration

District	Actual Flooding Rainfall (mm)		Rainfall Frequency (year)	
	1hr	3hr	1hr	3hr
Noksan-dong	43.0	64.5	2.4	1.8
Jeonmin-dong	59.6	82.6	5.5	2.9
Dunsan-dong	55.9	85.1	4.2	3.2
Hyomun-dong	39.5	96.0	3.0	7.3

**Table 12.** Model Improvement Effect

District	Actual Rainfall	Existing Model			Improvement Model			Decreasing Rate (%)
		Estimate (mm)	Error		Estimate (mm)	Error		
			(mm)	(%)		(mm)	(%)	
Noksan-dong	43.0 mm/1hr	52.3	+9.3	21.6	44.1	+1.1	2.6	19.1
Jeonmin-dong	59.6 mm/1hr	52.9	-6.7	11.2	60.2	+0.6	1.0	10.2
Dunsan-dong	55.9 mm/1hr	-	-	-	61.0	+5.1	9.1	-
Hyomun-dong	96.0 mm/3hr	95.3	-0.7	0.7	96.0	0.0	0.0	0.7

대전 전민동의 실제 침수강우량은 59.6 mm/hr이며, 기존 모델은 52.9 mm/hr, 개선모델은 60.2 mm/hr로 오차가 10.2% 감소하였다. 대전 둔산동은 55.9 mm/hr 강우에 의해 침수가 발생하였으며, 기존 모델은 침수위험기준을 예측하지 못하였으나, 개선 모델은 61.0 mm/hr을 예측하였으며 5.1 mm/hr의 오차가 발생하였다.

마지막으로 울산 효문동의 실제 침수강우량은 96.0 mm/3hr로 기존 모델은 95.3 mm/3hr, 개선 모델은 실제 침수강우량과 동일한 96.0 mm/3hr을 예측하여 0.7% 개선되었다.

따라서 3개 검토지역 모두 기존 모델 보다 개선 모델이 정확한 것으로 나타났으며, 1개 지역은 기존 모델에서 예측하지 못하였으나 개선된 모델에서는 예측 가능한 것으로 나타났다.

## 7. 결론

본 연구에서는 선행연구에서 제시된 유역특성기반의 침수 위험기준 예측 모델을 개선하기 위해 초기화 함수, 학습자료

의 수, 데이터 전처리를 통해 개선효과를 분석하여 최적 모델을 제시하였다. 또한 피해이력기반의 침수위험기준 추정이 불가능한 지역에 대한 침수위험기준을 예측하였으며, 실제 침수강우량과 기존 모델 및 개선 모델의 침수위험기준 예측결과를 비교하여 개선효과를 분석하고 적용성을 검증하였다.

Neuro-Fuzzy 알고리즘의 초기화 함수에 따른 개선효과를 분석하기 위해 선행연구에서 사용한 Grid Partition과 Subtractive Clustering기법을 모델 학습자료의 개수별로 비교하였다. Grid Partition의 경우 학습자료의 수가 적을 경우 정확성이 높은 것으로 나타났으나, 모델에 부정적인 영향을 미치는 학습자료가 포함된 경우 오차가 커지는 것으로 나타났다. 반면에 Subtractive Clustering 방법을 사용한 모델의 경우 학습자료가 증가함에 따라 우수한 성능을 보이고 있으며, 특히 학습자료의 군집화를 통해 부정적인 영향을 미치는 학습자료가 포함되어 있는 경우에도 비교적 정확성이 높은 것으로 나타났다.

본 연구에서 분석된 결과를 기존 모델(Model A) 대비 개선효과를 종합하여 Table 13에 나타내었다.

**Table 13.** Improvement Effect by Model

Division		Duration		Decreasing Rate (%)	
		1hr	3hr	1hr	3hr
Model A (Grid Partition, 27)	ERROR (mm)	15.8	36.1	-	-
	RMSE (mm)	20.9	45.5	-	-
Model H (Subtractive Clustering, 188)	ERROR (mm)	9.7	19.8	38.6	45.2
	RMSE (mm)	12.8	29.9	38.8	34.3
Model I (Subtractive Clustering, RanSAC)	ERROR (mm)	8.2	12.5	48.1	65.4
	RMSE (mm)	10.3	18.2	50.7	60.1

침수위험기준 추정 모델의 적용범위 확장 및 정확성 개선을 위해 학습자료를 추가로 구축하였으며, 학습자료의 수에 따른 모델 정확성을 비교하였다. 초기화 함수를 Subtractive Clustering, 188개의 학습자료를 사용한 모델은 지속시간 1시간에서 평균오차가 최대 38.6%, 3시간에서 45.2% 감소하는 것을 확인하였으며, RMSE도 각각 38.8%, 34.3% 감소하였다. 학습자료의 수가 증가하면서 모델의 정확도가 개선되는 것을 확인하였으며, 유역특성자료의 적용 범위도 증가하여 다양한 지역에 적용할 수 있을 것으로 판단된다.

마지막으로 데이터의 전처리 즉, 모델 학습과정에서 부정적인 영향을 주는 학습자료를 제거하기 위해 RanSAC 알고리즘을 적용하였다. 지속시간별로 정상분포를 벗어난 학습자료를 제외하여 모델을 구축한 결과 평균오차는 지속시간 1시간에서 48.1%, 3시간 64.5% 감소하는 것으로 나타났으며, RMSE도 각각 50.7%, 60.1% 감소하여 개선모델의 정확도는 증가하는 것으로 나타났다.

초기화 함수로 Subtractive Clustering을 적용하고, 188개의 학습자료에 RanSAC 기법을 통해 이상치를 제거하여 학습한 Model I의 정확도가 높은 것으로 나타났다. 따라서 Model I를 활용하여 피해이력이 부족하여 침수위험기준 추정이 불가능한 지역의 침수위험기준을 예측하였다. 예측된 침수위험기준의 개선효과 및 적용성을 확인하기 위해 실제 침수가 발생한 4개 지역을 대상으로 실제 침수발생강우량을 산정 하였으며, 기존모델과 개선모델을 비교한 결과 0.7%~19.1% 개선되는 것으로 나타났다.

본 연구에서 제시된 최적 모델(Model I)을 활용하여 피해이력이 없는 지역의 침수위험기준을 예측하고, 도시침수 예·경보를 위한 기준으로 사용한다면 도시침수로 인한 인명 및 재산 피해저감에 기여할 수 있을 것으로 판단된다. 또한 향후 실제 침수가 발생한 지역을 대상으로 침수위험기준 예측 결과와 실제침수발생 강우량 비교를 통해 지속적으로 개선한다면 정확한 도시침수 예·경보가 가능할 것으로 판단된다.

## References

- Centre for Research on th Epidemiology of Disasters (CRED). (2019). *Natural disasters 2018*.
- Cho, J.W., Bae, C.Y., and Kang, H.S. (2018). Development and application of urban flood alert criteria considering damage records and runoff characteristics. *Journal of Korea Water Resources Association*, Vol. 51, No. 1, pp. 1-10.
- Ministry of the Interior an Safety (MOIS). (2018). *Statistical yearbook of natural disaster 2017*.
- National Disaster Management Research Institute (NDMI). (2014). *A development of regional major disasters response scenarios and standards (I): Focused on urban flooding and drought*.
- National Disaster Management Research Institute (NDMI). (2015). *Establishment of foundation for regional urban flood response system (I)*.
- National Disaster Management Research Institute (NDMI). (2016). *Establishment of foundation for regional urban flood response system (II)*.
- National Disaster Management Research Institute (NDMI). (2017). *Development of technology on natural disaster risk assessment and urban flood forecasting*.
- Yeom, C.U., and Kwak, K.C. (2018). Performance comparison of ANFIS models by imput space partitioning methods. *Symmetry*, Vol. 10, No. 12, 700. doi:10.3390/sym10120700

---

Received	October 17, 2019
Revised	October 22, 2019
Accepted	November 11, 2019