

# ANFIS를 이용한 하천수위 예측

## Forecast of Stream Level Using ANFIS

최창원\* / 이재웅\*\*  
Choi, Changwon / Yi, Jaeeung

### 요지

최근 지구온난화로 인한 이상기후의 영향으로 강우일수는 줄고 있으나 강수량은 예년과 비슷한 수준을 보이고 있다. 이로 인해 갈수기의 용수부족 현상은 더욱 심해지고, 장마철의 홍수피해와 계릴라성 집중호우로 인한 피해가 커지는 등 해가 갈수록 홍수 예경보의 중요성은 더욱 높아지고 있다. 그럼에도 불구하고 현재 홍수 예경보 체계는 몇 가지 문제를 가지고 있다. 기존 예경보 체계의 경우 한 번의 예측을 수행하기 위해 수반되는 전처리과정과 주계산과정을 거치는 동안 각 과정에서 발생한 오차들이 반복, 누적되어 최종 결과물(예측된 유출량) 속에 모두 포함된다. 또한 기존 체계에서는 유출모형을 적용하기 위해서 토양형, 피복상태 등에 관련된 매개변수들이 필요한데, 이러한 매개변수의 결정에 어려움이 있고, 불확실성을 갖고 있다. 본 연구에서는 불확실성을 적극적으로 인정하고 수학적으로 해석하려는 fuzzy 이론을 신경망 이론에 도입하여 홍수 예경보 시스템의 운영과정에서 발생하는 불확실성의 문제를 해결하고자 하였다. 본 연구에서 사용한 ANFIS(Adaptive Neuro-Fuzzy Inference System)은 data driven model(자료에 기반을 둔 모형)의 하나로 다음과 같은 장점을 가진다. 우선 data driven model은 유역의 물리적, 지형적 특성을 고려하지 않고(매개변수 설정에서 발생하는 문제 해결 가능), 입력자료와 출력자료만을 고려하여 구축되는 모형이므로, 유역의 물리적 자료나 지형 자료와 같은 방대한 양의 자료 수집이 필요 없고, 일단 모형이 구축되면 자료의 입력만으로도 신뢰성 높은 결과를 단시간 내에 효율적으로 획득할 수 있다. 그리고 유역 내의 상황이 변화하더라도, 이들의 영향을 고려하여 쉽게 모형을 개선할 수 있다. 마지막으로 모형의 구축 과정이 물리적 모형에 비해 비교적 간편하다는 장점이 있다. 본 연구에서는 ANFIS를 통해 탄천유역의 강수량 자료와 대곡교의 수위자료를 입력자료로 사용하여 대곡교의 수위를 예측하였다. 입력 자료는 시간차 계열의 강우량과 수위 자료를 사용하였으며, 모형을 통하여  $t+1$ ,  $t+2$ ,  $t+3$  시간 후의 수위를 예측하였다.

핵심용어 : ANFIS, 신경회로망, data driven model, 수위 예측

### 1. 서론

물 관리의 영역에 있어서 치수는 인류를 홍수의 위험으로부터 생명을 지키고 생활의 공간을 마련해 준다는 점에서, 이수나 친수와 더불어 그 중요성이 높게 평가되어왔다. 우리나라의 경우 대부분의 하천은 유역면적이 작고 유로연장이 짧을 뿐만 아니라 산지가 많기 때문에 하천의 경사가 급한 곳이 많다. 또, 강우량의 대부분이 집중호우나 태풍 등의 형태로 편중되어 나타나기 때문에 치수의 중요성이 더욱 부각 되어왔다. 최근 들어 그 심각성을 더하고 있는 지구 온난화에 따른 현상들을 보면 치수의 중요성을 더욱 실감할 수 있다. 특히 온난화로 인한 현상들 중 계릴라성 집중호우의 피해는 실로 심각하다.

지금까지 사용해 왔던 홍수 예경보 과정에서는 특정 지점에서의 유출량을 예측하기 위해서 전처리과정과 주 계산과정을 거치는 동안 많은 오차들이 발생하고 그것들이 누적되어 예측치 속에 오차들이 내포되어 있

\* 아주대학교 대학원 건설교통공학과 석사과정

Research Assistant, Dept. of Civil & Transportation Engrg., Ajou Univ., Suwon 442-749, Korea  
(itsme99@ajou.ac.kr)

\*\* 아주대학교 공과대학 건설교통공학과 부교수

Associate Prof., Dept. of Civil & Transportation Engrg., Ajou Univ., Suwon 422-749, Korea (jeyi@ajou.ac.kr)

다. 또한 모형의 운영에 필요한 매개변수의 결정에서도 어려움이 있고, 많은 불확실성을 포함하고 있다. 최근 이러한 기준의 예경보 시스템에서의 한계를 극복하고 불확실성을 해결하여 모형의 정확도를 높이기 위한 방법으로, 불확실성을 적극적으로 인정하고 수학적으로 해석하려는 fuzzy 이론에 신경망을 도입한 neuro-fuzzy 이론이 활발히 연구되고 있다.

본 연구에서는 1965년 Zadeh가 처음 제안한 fuzzy 알고리즘과 Mamdani(1975)가 제어에 응용한 fuzzy 제어가 기초가 된 ANFIS(Adaptive Neuro-Fuzzy Inference System)를 이용하여 탄천 유역의 대곡교 수위를 예측하였다. ANFIS는 data driven model의 한 형태로 기존의 연구에 많이 사용되었던 물리적 모형과 달리 본 연구에서 사용한 data driven model은 다음과 같은 장점을 가지고 있다. 우선 data driven model은 유역의 물리적, 지형적 특성을 고려하지 않고, 강우와 같은 입력자료와 유출과 같은 출력자료만을 고려하여 구축되는 모형이므로, 유역의 물리적, 지형 자료와 같은 방대한 양의 자료 수집이 필요 없고, 일단, 모형이 구축되면 자료의 입력만으로 신뢰성 높은 결과를 단시간 내에 효율적으로 획득할 수 있다. 그리고 유역 내의 상황이 변화하면, 이들의 영향을 고려하여 쉽게 모형을 갱신할 수 있다. 마지막으로 모형의 구축 과정이 물리적 모형에 비해 비교적 간편하다는 장점이 있다.

## 2. ANFIS의 기본 이론

ANFIS는 신경회로망과 fuzzy 이론을 동시에 가장 쉽게 이용하는 방법으로 신경회로망의 구조와 학습능력을 이용하여 제어환경으로부터 얻은 입·출력 정보로부터 언어변수의 membership 함수와 제어규칙을 제어 대상에 맞게 자동 조종하는 방식을 사용한다.

효율적인 수자원 관리를 가장 어렵게 하는 것은 수자원관리에 필요한 많은 제약조건들에 불확실성이 존재한다는 것이다. ANFIS에서는 membership의 경계가 명확하지 않은 경우가 대부분인 현실의 문제를 다루기 위해서 fuzzy 집합과 fuzzy 추론의 개념을 도입한다. fuzzy추론에서는 단일fuzzy추론에서의 몇 가지 전제들이 교집합이나 합집합의 관계로 묶여있고 이를 통해 결론을 얻는 경우가 대부분이다. 각각의 fuzzy추론으로부터 얻은 결과를 종합하기 위해서는 각 결과값을 종합하여 도심을 구하는 방법으로 최종 결과값을 출력하게 된다. 이 과정을 탈fuzzy화(defuzzify)라 한다. fuzzy추론의 과정을 정리해보면 다음과 같다(table 1).

모형을 구축하기 위해 입력하는 강우, 유출자료 간의 상관성을 파악하기 위해 역전파 신경회로망 이론이 사용된다. 역전파 신경회로망은 Fig. 1과 같이 다층 퍼셉트론(multi-layered perceptron)의 구조를 갖고 있으며, 학습 단계와 산출 단계로 동작한다. 학습 단계에서는 입력 패턴과 목적 패턴이 다수의 입·출력 쌍으로 주어지고, 각 입력 패턴에 대해 출력을 먼저 계산(전방향 진행)한 후, 의도하는 출력과 실제 출력과의 차이를 줄이는 방향으로 연결 강도를 조정(역방향 진행)한다. 조정 과정이 끝나면 같은 과정으로 학습을 반복함으로써 최적의 연결 강도를 구한다. 산출 단계에서는 입력만 주어지면 연결 강도와의 계산에 의해 적절한 출력이 계산된다.

Table 1. process of fuzzy inference

단계	내용	결 차
1 단계	입력값의 fuzzy화	
2 단계	fuzzy연산자의 적용	
3 단계	합의법의 적용	
4 단계	모든 출력값의 종합	
5 단계	탈fuzzy화	

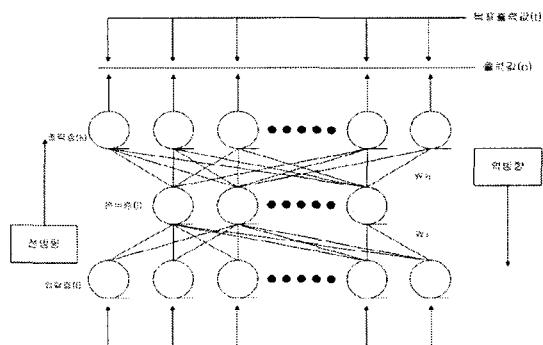


Fig. 1. structure of multi-layered perceptron

### 3. 모형의 적용

#### 3.1 입력 자료의 선정 및 구성

본 연구를 수행하는 과정에서, 최초의 연구는 구미교 수위관측소와 하대원동과 서현동의 두 개의 강우관측소의 자료를 이용하여 수행하였고, 이후 2006년 7월 15일을 기준으로 관측소를 변경·추가하여 최종 목표지점의 수위 예측을 실시하였으나 ANFIS을 통해 모델을 구축하고 검증하기 위해서 필요한 만큼의 강우사상이 발생하지 않았기 때문에 기존의 관측소를 기준으로 수행한 연구를 포함하였다(Fig. 2).

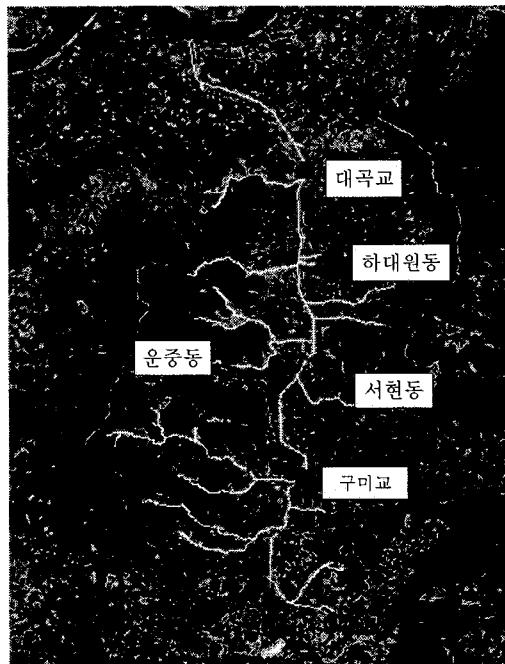


Fig. 2. river basin map of Tancheon

Table 2. mixing of all data set

강우관측소 (2 곳)	코드구성 (6 가지)	lag time (2 가지)	lead time (3 가지)
하대원동	P01 -Q01	10 min	$t+1$
	P01 -Q012		
서현동	P012 -Q01	30 min	$t+2$
	P012 -Q012		
	P0123-Q01		$t+3$
	P0123-Q012		

본 연구에 사용된 강우자료는 하대원동 강우관측소의 경우 2003년부터 2005년까지, 서현동 강우관측소의 경우는 2004년부터 2005년까지의 강우자료 중, 단일 호우사상이 100mm 이상의 강우량을 가지면서, 가능한 24시간 이상의 강우시간을 가지는 강우를 각각 세 개를 선정하였다. 두 개의 강우관측소에서 얻어진 자료들은 각각 별개로 실험을 수행하여 비교하였다.

수위예측 모델에 있어서 중요한 인자들 중 하나는 수위에 영향을 미치는 강우와 수위의 시간적 분포를 입력자료 구성에 어떻게, 얼마만큼 배치하는 것이다. 수위와 강우자료를 어떻게 구성하는 것이 미래의 수위를 가장 우수하게 예측하는지를 비교 분석하는 것이 본 연구의 수행에서 가장 중요한 부분 중의 하나이다. 본 연구에 있어서는 어느 입력 자료 구성이 가장 좋은지를 찾기 위해 두 곳의 강우관측소에 대해, 6가지 강우-수위 자료조합, 2가지의 lag time, 3가지의 lead time을 설정 조합하여 각각의 강우 관측소 자료마다 36가지의 경우에 대해 수위를 예측하고 비교하였다(Table 2).

#### 3.2 최적 자료 구성과 예측 결과 분석

본 연구에서는 최적의 자료 구성을 찾기 위해서 RMSE(Root Mean Square Error), 첨두수위비, 그리고 예측형상의 세가지를 비교하여 종합적으로 판단하고 최적의 구성을 선정하였다. 먼저, 관측 자료와 예측값의 비교를 통해 우수한 예측이 된 모형을 찾을 때는 전체적인 추세와, 상승부에서의 정확도 특히 첨두수위에서의 정확도에 중점을 두었고, 또한 첨두값에서 실측 수위보다 낮은 값을 예측하는 것보다는 높은 값을 예측 하는 것

에 중점을 두었다. 본 연구의 목표인 홍수 예경보 시스템의 구축에서는 첨두수위의 정확한 예측이 가장 중요한 요소 중의 하나이다. 따라서 RMSE 값을 통해서는 전체적인 추세만을 판단하고, 첨두수위비를 통해서는 어떤 자료 구성이 가장 적합한지를 찾아내는 방법이 옳다고 판단된다.

결국 예측형상, RMSE, 첨두수위비를 통해서 자료의 구성을 비교해보았을 때,  $t$ ,  $t-1$ ,  $t-2$  시간의 강우자료와  $t$ ,  $t-1$  시간의 수위자료를 사용하여 자료를 구성했을 때 가장 우수한 수위 예측을 하는 것으로 판단된다.

### 3.3 변경된 관측소를 이용한 대곡교 수위 예측

변경된 관측소의 자료로 모형을 구성하기 위해서 세 지점(하대원동, 운중동, 서현동)의 강우자료와 한 지점(대곡교)의 수위자료를 사용하였다(Fig. 2 참조). 탄천 유역은 세 곳 관측소의 강우값을 각각 서로 다른 입력변수로 고려할 만큼 지체시간이 길지 않다고 판단하여 유역 전체를 대표하는 하나의 강우량 값으로 평균한 값을 사용하였다. 유역의 강우량은 세 곳의 강우관측소에서 측정 된 값을 티센법을 이용하여 유역전체의 평균값으로 사용하였다.

앞에서 언급한 것처럼 관측소를 변경·추가한 후에는 두 번의 강우사상만이 발생하여서 기존에 선정된 입력자료 구성을 기존의 관측 데이터를 사용하였고, 관측소 변경 후 실제 수위예측을 시행할 때 ANFIS가 원활히 작동하는지 확인을 위해서 checking 자료 중 일부를 testing 자료로 사용하여 시험 적용하였다. testing 자료의 경우 실제 연구의 적용에서 강우가 발생하여 수위가 상승하고 있는 시점에서 수위를 예측하게 되므로 전체적으로 수위가 상승하는 형태를 갖는 부분을 testing 자료로 선정하였다.

### 3.4 결과

우선 testing 자료가 checking 자료의 일부이기 때문에 수위예측 모형이 만들어지는 과정에서 이미 매개 변수 추정에 사용된 자료임을 고려하지 않고 별개의 testing 자료라는 가정하에 판단하였다. 결과를 분석하면, RMSE값과 첨두수위비 모두 우수한 결과를 보였다. 예측값은 실제 측정치와 거의 비슷한 값으로 예측되었고(Fig. 3), 실측값과 예측값의 차이가 3cm 이하의 양호한 값을 나타냈다(Table 3).

그리고 이번 예측의 경우 선행시간에 따른 RMSE값을 비교해보았을 때 예측시간이  $t+1$ 인 경우가  $t+2$ 의 경우보다,  $t+2$ 의 경우가  $t+3$ 인 경우보다 우수하다는 것을 알 수 있었다. testing 자료가 전체적으로 상승하는 형태를 띠고 있기 때문에 첨두수위값 자체가 측정 데이터의 마지막 관측값에서  $t+1$ ,  $t+2$ ,  $t+3$ 으로 갈수록 증가하고 있고, 각각의 마지막 데이터가 첨두수위를 갖고 있기 때문에 첨두수위비의 경우 서로 비교하는 것이 무의미하다고 판단된다.

Table 3. RMSE, peak ratio and peak error

	RMSE	첨두수위비 (%)	첨두값 오차(m)
$t+1$	0.026152	0.103054	-0.02459
$t+2$	0.031655	0.351987	-0.01434
$t+3$	0.043210	0.001796	-0.02992

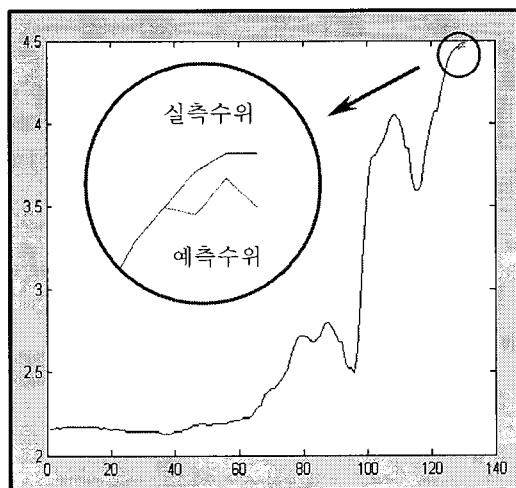


Fig. 3. forecast of stream level

## 4. 결 론

본 연구에서는 탄천유역의 대곡교 지점의 수위를 예측하기 위해 fuzzy이론과 신경망 이론이 결합된 neuro-fuzzy 시스템인 ANFIS를 이용하여 수위 예측 모형을 구성하고, 비교·검토 하였다. ANFIS를 이용하여 모형을 구성하고, 비교하기 위해서는 최소 세 번 이상의 강우사상이 필요하지만, 관측소를 변경·추가한 후 모형에 사용 가능한 강우사상은 두 번밖에 발생하지 않았다. 그래서 이번 연구에서는 기존의 관측소의 자료를 이용해서 수위 예측 모델을 구성하고, 비교·검토 하여 가장 우수한 예측모델을 생성하는 입력 자료를 구성하였고, 관측소 변경·추가 후의 자료에 대입하는 것으로 연구를 진행하였다. 이를 통해 다음과 같은 결론을 얻을 수가 있었다.

1) ANFIS 모형을 구축하는데 있어서 입력 자료의 개수가 많은 것이 좋은 모형을 구축하게 하는 조건이 되지 못하고, 관측 자료의 특성에 맞게 입력 자료가 조합되었을 때 가장 좋은 예측 모형을 구축할 수 있다. 그러나 자료로 사용하는 각각의 강우 사상의 자료 길이가 긴 것은 좀 더 좋은 모형을 구축하거나, 구축 후에 좀 더 우수한 예측을 하는데 도움이 된다고 판단된다.

2) 수위를 예측하는데 있어서 수위의 변화 양상 특히, 상승부와 하강부의 형상에 따라서 정확도의 차이가 나타나는 것으로 보인다. 상승부나 하강부의 형상이 완만한 경사를 가지고 천천히 증감하는 형태를 보일 때 예측값이 실측치와 비슷한 값을 가졌다. 반면, 수위의 형상이 가파른 경사를 가지거나 갑작스런 상승이나 하강을 할 경우 완만한 경사를 갖는 경우보다 같은 변동폭에 대해 적은 데이터를 가지기 때문에 더 큰 편차를 보이는 것으로 사료된다.

3) 기존의 관측소를 이용한 수위 예측에서는 세 개의 시계열 강우자료와 두 개의 시계열 수위자료의 형태로 입력 자료가 구성되었을 때 가장 우수한 예측을 하였다. 변경, 증설된 관측소를 이용한 수위 예측의 경우에는 두 번의 강우사상만이 발생하여 제대로 완성된 형태의 모델을 구성하고, 비교 검토할 수는 없었지만, 기존의 관측소를 이용하여 완성한 입력 자료 구성을 이용하여 testing해 봤을 때는 어느 정도 양호하게 예측을 한 것으로 판단된다.

4) 마지막으로 이번 연구를 수행함에 있어서 강우사상의 부족으로 연구가 완벽하게 마무리되지 못한 것 같은 아쉬움은 있지만, 앞으로 몇 차례의 추가 강우사상이 발생하면, 본 연구를 통해 얻은 결론을 바탕으로 training, checking, testing 자료를 뚜렷한 수위 변동양상을 나타내고, 강우 시간이 수위 예측을 잘 하기에 충분히 긴 강우사상을 선택 하면 지금까지의 연구 결과에서 얻은 결과보다 더 좋은 수위 예상 결과를 얻을 수 있을 것이다.

## 참 고 문 헌

- Mamdani, E.H., and Assilian, S. (1975). "An experiment in linguistic synthesis with a fuzzy logic controller", *International Journal of Man-Machine Studies*, Vol. 7, No. 1, pp. 1-13.  
Zadeh L.A. (1965). "Fuzzy Sets", *Information and Control*, Vol. 8, pp.338-353.