

인공신경망을 이용한 RDAPS 강수량 예측 정확도 향상

Improving Accuracy of RDAPS Prediction Precipitation using Artificial Neural Networks

신주영*, 최지안**, 정창삼***, 허준행****

Ju-Young Shin, Gian Choi, Chang-Sam Jeong, Jun-Haeng Heo

요 지

이 연구는 기상수치예보 모델 중 지역수치예보모델인 RDAPS 모델을 이용하여 강우자료를 예측한 값과 실제 강우관측지점에서의 강우량을 비교해 보고 RDAPS 예측량의 정확도를 높이기 위한 연구이다. RDAPS 모델의 자료는 00UTC와 12UTC에 3시간 누적 자료를 48시간에 대해서 생성하고, 30km 격자망에 대한 정보를 담고 있기 때문에 1시간 간격으로 측정된 지점 강우량과의 비교를 위해서는 관측지점과 근거리 정보를 찾고 1시간 간격의 관측 자료를 3시간 누적강우량으로 바꾸는 전처리 과정이 필요하다.

실제 강우예측이 어려움을 겪는 것처럼 RDAPS의 예측 강우량과 관측 강우량은 큰 차이를 보이는 것으로 나타났다. 예측 강우량의 정확도를 높이고자 인공신경망을 적용하였다. 인공신경망이란 뇌기능의 특성 몇 가지를 컴퓨터 시뮬레이션으로 표현하는 것을 목표로 하는 수학 모델이다. 강우수치예측 자료 외에도 RDAPS 모델에서 얻을 수 있는 풍향, 풍속, 상대습도, 기압, 온도 등의 다른 수치자료들을 이용하여 인공신경망을 이용하여 자료들의 패턴을 시뮬레이션 하여 정확도가 높은 예측값을 얻을 수 있었다.

핵심용어: RDAPS, 인공신경망, 강우량 수치예보

1. 서론

최근 전 세계적으로 국가적 재난 형태의 기상재해가 빈발하고 있다. 자연재해에 의한 피해예방뿐만 아니라 국민의 삶의 질 향상, 지속가능한 발전 등 다양한 분야에서 기상의 중요성이 증가하고 있다. 우리나라의 경우 각종 기상정보 중 강우 예보의 정확성 미흡으로 작게는 개인적인 손실에서부터 크게는 국가적 손실에 이르는 피해가 발생하고 있다. 이러한 피해를 감소시키고, 급변하는 기상현상 변화에 대비하기 위해서는 정확한 기상예보가 반드시 필요하다.

우리나라는 단기적인 수자원 운영에 대한 계획 수립과 장기적인 수자원확보 및 활용을 위해서 RDAPS(Regional Data Assimilation and Prediction System) 기상수치예보 모델을 이용하여 수치예보를 하고 있다. RDAPS 모델의 자료는 00UTC와 12UTC에 3시간 누적 자료를 48시간에 대해서 생성하고, 30km 격자망에 대한 정보를 담고 있기 때문에 1시간 간격으로 측정된 지점 강우량과의 비교를 위해서는 관측지점과 근거리 정보를 찾고 1시간 간격의 관측 자료를 3시간 누적강우량으로 바꾸는 전처리 과정이 필요하다. 전처리한 지점강우량과 RDAPS 모델의 자료는 비교적 큰 오차를 보여 기상현상의 정확한 예보가 얼마나 어려운지를 보여준다.

* 정희원 · 연세대학교 대학원 토목공학과 석사과정 · E-mail: ausran@yonsei.ac.kr

** 정희원 · 연세대학교 대학원 토목공학과 석사과정 · E-mail: gian@yonsei.ac.kr

*** 정희원 · 인덕대학 건설환경설계학과 전임강사 · E-mail: csjeong@induk.ac.kr

**** 정희원 · 연세대학교 사회환경시스템공학부 토목환경공학전공 교수 · E-mail: jhheo@yonsei.ac.kr

자료의 오차를 보완하기 위하여 수치 예보값에 대한 통계역학적 방법이 사용된다. 주로 많이 쓰이는 방법은 과거강수자료를 사용하여 강수와 상관성이 높은 기상변수들을 추출하여 다중회귀방정식을 개발한 후 이를 이용하여 강수확률모형을 만드는 PPM(Perfect Prog Method)와 통계상관모형을 구성할 때 과거 수치예보까지를 고려하여 모형을 구성하는 MOS(Model Output Statistics) 방법이 있다. 우리나라 기상청에는 PPM을 이용한 18개 구역에 대한 12시간 간격의 48시간까지의 강수확률이 예보되고 있다. 본 연구에서는 MOS(Model Output Statistics)를 적용하여 RDAPS 시계열을 개선하였다. RDAPS 모델의 강수량 자료의 정확도를 높이기 위해서 인공신경망을 적용하였고, 강수수치예측 자료 외에도 RDAPS 모델에서 얻을 수 있는 풍향, 풍속, 상대습도, 기압, 온도 등의 다른 수치예보자료들을 인공신경망에 학습시켜 더 높은 정확도의 예측 값을 얻을 수 있었다.

2. 기상 수치예보 자료 및 지점강수량 자료

2.1 RDAPS(Regional Data Assimilation and Prediction System)

기상청 지역예보 모델로 사용하는 RDAPS는 동아시아영역 예보자료로 예보 영역은 30km 격자간격으로, 191 × 171 개의 격자수로 이루어져 있다. 좌표값 (1,1)에 해당하는 격자점의 위도는 12.269°N이고, 경도는 102.477°E 이다. 연직층수는 24개로 1,000, 975, 950, 925, 900, 875, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 200, 150, 100, 70, 50mb 면으로 구성되어 있다. 자료의 생산은 하루 2번으로 00UTC와 12UTC이며, 3시간 간격의 예보자료로 총 48시간 예보자료를 포함하고 있다. 표 1은 자료의 변수들에 대한 정보이다(한국수자원공사, 2005).

표 1. RDAPS 자료의 변수

변수명	설명	단위
APCPsfc	누적 강수량	kg/m ²
HGTsfc	지표 지오폠펜설고도	gpm
HGTprs	지오폠펜설고도	gpm
MIXRsfc	지표 혼합비	kg/kg
MIXRprs	혼합비	kg/kg
PRMSmsl	해면기압	Pa
TMPsfc	지표 기온	K
TMPprs	기온	K
UGRDsfc	지표 동서방향 바람	m/s
UGRDprs	동서방향 바람	m/s
VGRDsfc	지표 남북방향 바람	m/s
VGRDprs	남북방향 바람	m/s

표 2. 지점강우관측소 정보

번호	관측소명	TM_X	TM_Y
1	충주	284756.6542	385849.0508
2	청주	239342.1991	348768.2730
3	대전	233325.0650	319079.5606
4	추풍령	289334.4035	302609.0725
5	군산	173393.2224	277007.9193
6	전주	213933.6652	257954.7184
7	이천	242881.6265	418130.1166
8	보은	265711.2063	332109.7061
9	천안	210565.9010	364252.1276
10	보령	160186.9744	314135.7447
11	부여	192821.1268	307962.6012
12	금산	243300.6287	289568.9743
13	장수	247066.5310	239815.5291
14	거창	282421.9600	241715.3700

2.2 지점강수량 자료

이 연구의 대상 지역으로는 위치상으로 우리나라의 중앙부에 위치하여 우리나라의 평균적인 강우 패턴을 보이는 금강유역을 선정하였다. 이러한 특징 때문에 금강유역을 선정하였다. 금강유역중에서도 신뢰성 있는 자료를 보유하고 있는 기상청 지점을 이용하였고, 그중에서도 RDAPS 모델지점과 가장 근접한 지점을 선정하여 자료를 비교하였다. 표 2는 연구에서 사용된 14개 지점과 좌표값을 나타내고 있다.

3. 인공신경망을 이용한 강수량 보정

3.1 인공신경망 이론

인공신경망이란 시냅스의 결합으로 네트워크를 형성한 인공 뉴런(노드)이 학습을 통해 시냅스의 결합 세기를 변화시켜, 문제 해결 능력을 가지는 모델 전반을 가리킨다. 신경망 모형에서 학습이란 입력층, 은닉층, 출력층으로 구성된 다층신경망의 각 층 노드들의 연결강도를 최적의 상태로 적응시키는 과정을 말하며 전반적인 개형은 그림 2와 같다. 단, 은닉층은 1개 이상으로 할 수 있다. 인공신경망은 수학적 연산능력을 가진 노드들이 상호 연결된 것으로 적절한 학습규칙에 의하여 작동한다. 각 노드들은 결합함수(식 (1))와 전달함수(활성화 함수)(식 (2))에 의해 수학적 연산을 수행한다(C. M. Bishop, 1995).

$$S_j = \sum_{i=0}^m w_{ji}x_i \quad (1)$$

$$O_k = \frac{1}{1 + \exp(-S_j)} \quad (2)$$

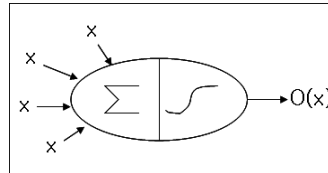


그림 1. 노드에서의 함수

여기서 w_{ji} 은 각 층으로 연결된 연결강도이며, x_i 는 입력층의 입력값이다.

그림 1은 실제 노드로 입력되는 신호는 가중치가 부여된 값들의 합과 전달함수를 나타낸 그림으로서 식 (1)과 같이 나타낼 수 있다(박성천, 2006). 입력신호에 가중치가 부여된 값들의 합은 시그모이드 함수(비선형함수)를 통해 노드에서의 출력값을 도출해낸다. 연결강도를 최적화하기 위해 사용된 학습 알고리즘은 오차역전파등의 학습 알고리즘을 적용한다.

3.2 입력변수 결정

RDAPS 에서는 총 152개의 자료가 나온다. 152개의 자료를 모두 인공신경망의 입력층에 넣어서 학습을 시킬 경우 프로그램이 복잡해지고, 학습시키는데 시간이 오래걸린다. 이럴 때는 강우가 내리는 데 밀접한 영향이 있는 인자와, 강우현상에 필수 요소들 위주로 입력값을 구성하면 더 좋은 결과를 얻을 수 있다. 본 연구에서는 152개의 주요인자중 4가지 변수를 선정하였다. 표 3은 사용된 변수들의 상관계수 값을 나타낸 표이다. 상관성 분석을 통해서 가장 상관성 계수가 높은 M850(850mb에서의 혼합비), 우리가 보정하고자 하는 값인 지표면 강우량값(ASFC), 그리고 강우이동을 가장 잘 나타내리라 판단되는 700mb의 풍향자료 U700(동서방향의 풍향)과 V700(남북방향의 풍향)을 이용하였다(김광섭, 2006).

표 3. 입력변수

변수명	상관계수
ASFC	0.045
M850	0.274
U700	-0.131
V700	0.058

3.3 모델구축

본 연구에 사용된 인공신경망의 종류는 다층(multi-layer)신경망으로 입력층 노드(node)는 4개로 구성되어 있으며, 각 노드(node)의 입력인자로는 UGRD 700bm, VGRF 700mb, MIXR 850mb, APCP이다. 은닉층은 1개 층으로 구성하였고, 은닉층은 총 4개의 노드(node)로 선정 하였으며, 출력층 노드(node)는 1개로 개선된 강수량 값을 출력하도록 구성하였다. 그림 2는 본 과제에서 적용한 인공신경망 모형을 나타낸 그림이다.

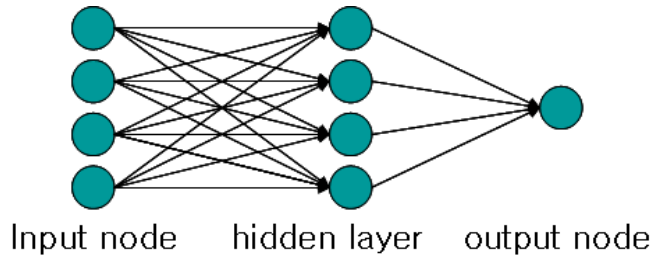


그림 2. 인공신경망 모형

4. 결과 분석

개선된 예측 시계열 자료를 단기 시계열에서 처음 면적평균강수량(MAP)를 계산하여 순수 RDAPS 강수량값과 관측 강수량을 비교하였다. 그림 3은 전체 지역의 2005년 총강수량과 각 소유역별 총강수량에 대하여 RDAPS 강우자료(초록색)와 관측자료(붉은색) 그리고 인공신경망 보정자료(파란색) 각각의 면적평균강수량(MAP)을 비교한 그림이다. 실측자료와 비교했을 때, RDAPS 강우자료보다 보정된 강우량이 실측자료와 유사하다는 것을 알 수 있다. 그림 4는 전체 지역의 2005년 총강수량과 각 소유역별 총강수량에 대한 관측 면적평균강수량(MAP)와 순수 RDAPS자료와 보정자료로 예측한 면적평균강수량(MAP)의 상대오차를 나타냈다. 그림 4를 보면 전체 지역의 연 총강수량의 상대오차는 RDAPS 자료가 -65%미만으로, 보정자료는 5%미만으로 나타났으며, 각 소유역별 연 총강수량의 상대오차는 RDAPS 자료가 -75 ~ -65%내외로, 보정자료는 5% ~ -15% 내외로 나타났다. 그림 5는 월별 총 관측 면적평균강수량과 월별 총 예측 면적평균강수량의 절대상대오차를 평균한 값을 나타낸 그래프이다. 그림 5를 보면 전체지역의 절대상대오차평균이 RDAPS 자료가 70%, 보정자료는 약 150%정도 나오고, 소유역별로는 RDAPS 자료의 경우에는 50% ~ 100%, 보정자료는 절대상대오차평균이 50% ~ 350%정도로 나타났다. 연 총강수량 비교와 연 총강수량 상대오차 비교에서는 인공신경망으로 보정한 강우자료가 순수 RDAPS 자료보다 상대오차도 적고, 실측값과 유사한 값이 나왔다. 하지만 월별 절대상대오차평균 비교시에는 RDAPS 자료보다 인공신경망으로 보정한 값이 더 오차가 큰 것으로 나타났다. 이는 인공신경망으로 보정한 자료는 연 총강수량적인 면에서는 높은 정확도를 보이거나 월별로 정확도를 계산하였을 때는 매우 낮은 정확도의 값이 나온다는 것을 의미한다.

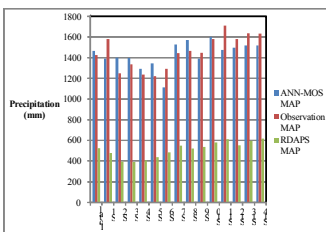


그림 3. 연 총강수량 비교

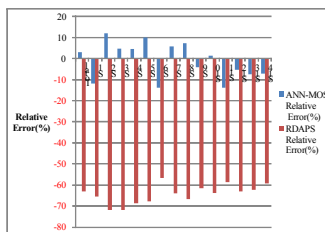


그림 4. 연 총강수량 상대오차 비교

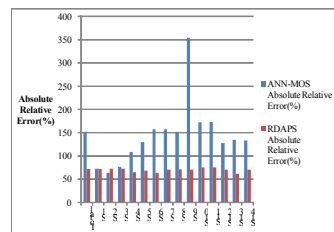


그림 5. 월별 절대상대오차평균 비교

5. 결론

본 연구에서는 금강유역의 2005년 RDAPS 기상수치모델 자료를 이용한 예측 강우산정량이 실측값과 차이가 크다는 것을 파악하고, 강우 예측값의 정확도를 높이기 위해서 인공신경망을 이용하여 강우자료를 보정하였다. 인공신경망은 입력층, 은닉층, 출력층으로 구성되어있고, 각 층은 연결강도로 연결 되어있다. 인공신경망은 연결강도 수정의 과정을 되풀이하여 적절한 연결강도를 찾아낸다. 입력층에 RDAPS에서 얻어진 자료를 넣고(UGRD 700bm, VGRF 700mb, MIXR 850mb, APCP) 출력값이 측정된 강우량과 비슷해지도록 연결강도를 조절하여 보정된 강우량 값을 얻을 수 있다. 인공신경망을 이용하여 RDAPS 기상수치모형에서 얻은 강우량 자료를 보정한 후 실측값과 비교해 보았을 때, 정확도가 증가하는 것을 볼 수 있다. 하지만 총량적인 예측의 정확도에 비하면 월별 예측은 정확도가 낮다는 것을 알 수 있다. 향후연구에는 월별 예측의 정확도가 낮은 이유와 그에 따른 해결 방법을 찾는 연구가 이어져야 할 것이다.

감사의 글

본 연구는 21세기 프론티어연구개발사업인 수자원의 지속적 확보기술개발사업단의 연구비지원(과제번호: 1-6-3)에 의해 수행되었습니다.

참고문헌

1. 김광섭 (2006) 상층기상자료와 신경망기법을 이용한 면적강우 예측 한국수자원학회논문집 제 39권 제8호 pp.717-726
2. 박성천, 오창열, 김동렬, 진영훈, (2006) 인공신경망 이론을 이용한 홍수유출 예측시스템 개발 -GULFFS 개발 및 적용- 대한토목학회논문집 제26권 제 2B호 pp.145-152
3. 한국수자원공사. (2005). 유역 물 관리 운영 기술 개발
4. Christopfer M. Bishop (1995) Neural Networks for Pattern Recognition OXFORD pp.140-148