

인공신경망모형을 이용한 대규모 대기모형모의결과의 댐유역스케일에서의 지역화기법

Dam Basin-scale Regionalization of Large-scale Model Output
using the Artificial Neural Network

강부식*·이봉기**

Kang, Boosik·Lee, Bongki

요 지

본 연구에서는 GCM 기후변화 전망 시나리오를 이용하여 유역단위의 기후변화를 추정하였다. 원시 GCM 시나리오를 지역화 시키기 위해서 인공신경망 모형을 사용하였다. GCM에서 모의되는 강수플럭스, 해면기압, 지표면 근처에서의 일 평균온도, 지표면으로부터 발생하는 잠열플럭스 등과 같은 22개의 변수는 인공신경망의 잠재적 예측인자로 사용되었으며, AWS에서 관측된 강수량과 온도는 예측변수로 사용되었다. 원시 GCM 데이터는 CCCma(Canadian Centre for Climate Modeling and Analysis)에서 제공되는 CGCM3.1/T63 20C3M 시나리오를 사용하였으며, 인공신경망 학습과정에서 사용된 기준시나리오(reference scenario)자료의 기간은 1997년부터 2000년까지의 데이터를 사용하였다. 인공신경망을 학습을 통하여 결정된 각 층사이의 가중치를 이용하여 이산화탄소 배출농도를 가정하여 생성된 CGCM3.1/T63 SRES B1 기후변화시나리오(project scenario)를 인공신경망의 입력값으로 하여 미래의 기온과 강수변화를 전망하였다. 신경망의 학습효과를 높이기 위하여 기온과 강수에 대한 평균 및 누적기간을 각각 일단위와 월단위로 설정하였다.

본 연구에서 사용된 인공신경망은 3층 퍼셉트론(다층 퍼셉트론)을 사용하였으며, 학습방법으로는 역전파알고리즘(back-propagation algorithm)을 이용하였다. 민감도분석을 통하여 선택된 예측인자는 소양강댐유역(1011, 1012소유역)에서의 인공신경망 예측인자로 활용되었으며, 2001년부터 2100년까지의 일 평균온도와 일 강수량의 변화경향을 추정하였다. 1011유역, 1012유역에서는 여름철의 온도변화경향이 겨울철에 비하여 높게 나타났다.

일 평균온도의 통계분석 결과 평균예측오차가 가장 적게 나타나는 지역은 1001유역으로 -0.08 로 평균예측오차가 가장 적게 나타났으며, 인공신경망기법을 이용하여 스케일 상세화된 일 평균온도와 관측된 일 평균온도가 얼마나 잘 일치하는지를 확인할 수 있는 1012유역에서 CORR이 0.74로 가장 높게 나타났다.

핵심용어: 인공신경망(Artificial Neural Network), GCM, 지역화(downscaling)

* 정회원·단국대학교 토목환경공학과 조교수·공학박사·031-8005-3484(E-mail: bskang@dankook.ac.kr)

** 정회원·(주)도화종합기술공사 사원·공학석사·02-3469-7358 (E-mail: leebk@dankook.ac.kr)

1. 서론

홍수관리와 안정적인 용수공급을 위해서는 하천의 유량예측을 위한 강수량의 예측은 수자원관리에 있어서 중요한 요소이다. 더 나아가 기후변화로 인한 강수량의 변화가 수자원에 미치는 영향을 파악하기 위해서는 강수량의 장기예측은 치수관리에 필요한 댐과 여수로와 같은 주요시설물 설계에 필요하다. IPCC(2007) 보고서에 의하면 기후계의 변동으로 인하여 지구의 평균 지표기온은 1906년부터 2005년까지 약 0.74℃ 상승하였으며, 강수의 경우 온도와 기상조건에 크게 영향을 받으며 온도가 1℃ 상승할 때 대기의 수분 보유력은 약 7% 증가한다고 설명하고 있다. 따라서 온도의 변화는 수증기의 증가를 가져올 것이며, 강수의 발생빈도가 더 높아질 것이다. 이러한 지역규모의 기후변화영향에 대한 연구수행을 위해 수학적 방정식을 이용하여 대기과정을 설명하는 대기대순환모형(GCMs; General Circulation Models)이 사용되고 있다(Tripathi, 2006). GCMs은 공간해상도(spatial resolution)가 작기 때문에 수자원분야에서 주로 사용되는 유역규모에서의 스케일특성과 물리적 특징을 고유적으로 설명할 수 없다. 따라서 GCMs에서 계산된 모의값을 수자원분야에서 사용하기 위해서는 강수의 시·공간적 특성을 고려하는 지역화기법(downscaling)이 사용되고 있다. 본 연구에서는 GCM이 고유적으로 가지고 있는 저해상도의 모의값을 유역규모 혹은 지역규모에서의 시공간적인 스케일특성을 부여함으로써 대상유역에서의 월누적강수량과 일평균온도를 장기적인 변화추세를 예측하는데 있다.

Wilby 등(1998)은 대기대순환모형의 모의값을 이용하여 수문기상학적 변수들을 통계적 스케일상세화를 실시하였다. Landman 등(2001)은 대상유역에서 12개 댐 유입부에서의 유입유량을 3가지 범주형태로 예측하기 위하여 통계적 스케일 상세화기법을 이용하였다. Dibike(2006)가 NCEP 재분석자료와 CGCM1를 이용하여 인공신경망기법을 이용하여 강수와 온도예측을 수행하였으며, Tripathi(2006)는 support vector machine기반의 다운스케일링을 통하여 강수를 예측하였다. Stehlik와 Bárdossy(2002)는 일반적인 강수의 순환패턴과는 다르게 새로운 퍼지규칙(fuzzy rule)을 기반으로 순환패턴을 분류하여 스케일 상세화의 입력값으로 사용하고 다른 기상조건을 가진 지역들에 성공적으로 적용하였으며, Wetterhall 등(2005)은 강수와 유출량과 같은 지역 고유의 매개변수는 스케일 상세화기법이 필요함을 설명 하였다. Valverde Ramirez(2005)는 브라질 상파울로지역에서 강우예측을 위해 인공신경망과 다중선형회귀를 비교·연구한 결과 인공신경망이 다중선형 회귀모형보다 신뢰성이 있음을 설명하였으며 인공신경망(ANN; Artificial Neural Network)은 장기적 강수예측에 있어서 좋은 보정기법이 될 수 있다고 설명하였다.

2. 연구의 이론적 배경 및 대상유역

2.1 인공신경망

인공신경망은 인간의 두뇌를 구성하고 있는 기본단위인 뉴런(neuron)의 개념을 도입한 이론을 기초로 뉴런의 인식과정을 수학적인 모형으로 일반화시키기 위해 개발된 병렬정보처리 시스템이다. 인공신경망은 신경망의구의 계층구조에 따라 입력층과 출력층만 존재하는 단층신경망과 입력층, 하나 이상의 중간층(은닉층), 출력층이 존재하는 다층신경망으로 구분된다. 여기서, 입력층이란 각각의 입력뉴런들이 입력되는 층을 말하며, 은닉층은 은닉뉴런이 존재하는 층, 출력층은 출력뉴런이 존재하는 층을 말한다. 일반적으로 하나 이상의 은닉층을 가지고 있는 다층신경망이 많이 사용되어오고 있으며, 다층신경망의 일반적인 구조는 그림 1과 같다. 각각의 뉴런들은 서로 연결되어 있는데 이것은 각각의 뉴런들이 생물학적인 뉴런에서의 시냅스의 역할을 하며, 인공신경망에서는 연결강도 혹은 가중치벡터라 일컫는다.

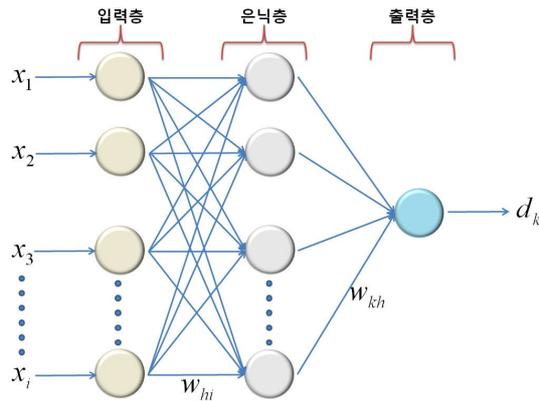


그림 1. 다층신경망의 구조

인공신경망모형은 입력뉴런, 은닉뉴런의 수, 학습할 케이스의 수가 학습성과에 중요한 영향을 미친다. 이는 은닉뉴런의 수가 많아지면 출력값과 목표값 사이의 오차를 계산하는 과정이 복잡해지고, 계산시간도 더 많이 소비하기 때문이다. 이와 더불어 학습할 케이스의 수가 많다면 차원의 저주(curse of dimensionality)에 빠지기 쉽다. 차원의 저주란 수학적 공간에 별도의 차원을 더할 때 총량이 기하급수적으로 늘어나는 문제를 말한다(Bellman, 1957). 이러한 문제를 해결하기 위해서 목표값에 관련성이 적거나 필요 없는 예측인자보다 목표값에 영향을 미치는 예측인자에는 더 많은 가중치를 부여하는 알고리즘들이 개발되어왔다.

본 연구에서는 차원의 저주를 피하기 위하여 예측인자가 목표값에 미치는 영향을 파악하기 위하여 민감도 분석을 실시하였다. 예측인자로 사용된 기후변화시나리오인 CGCM3.1/T63의 2D 변수들을 기준으로 입력변수를 하나씩 제거하면서 출력변수에 미치는 영향을 파악하였다. 출력값과 목표값 사이의 오차를 이용하여 가중치를 계산하는 역전파알고리즘(backpropagation algorithm)을 인공신경망모형의 학습방법으로 사용하였다.

2.2 역전파알고리즘

역전파알고리즘은 인공신경망의 학습방법중 하나로 일반적으로 많이 사용되어오고 있는 방법이다. 역전파알고리즘은 미분 가능한 활성화함수를 사용하여 경사하강법을 통하여 오차함수의 최소값을 찾아 가중치의 크기를 결정하는 방법이다(Rojas, 1996). 하지만 역전파알고리즘은 출력값과 목표값사이의 오차를 이용하여 오차함수를 편미분함으로써 가중치를 줄여나가는 방법을 통하여 학습을 수행하게 되는데 최적의 해인 전역 최소값(global minima)가 아닌 학습 데이터의 특성으로 인해 국소적으로 낮은 값을 을 전역 최소값(global maximum)으로 인식하게 되어 인공신경망의 수행성과를 저하시키는 문제를 안고 있다. 이를 피하기 위해서는 인공신경망 학습시 신경망의 학습속도 관련된 학습률을 작은 값으로 설정하여 점점 줄여나가고, 가중치 변화량에 영향을 미치는 모멘텀을 작게하여 국부최소값을 전역 최소값으로 인식하지 않게 하는 방법이 있다. 본 연에서는 식 1과 같은 오차함수를 사용하였으며, 은닉층(혹은 입력층)과 출력층(혹은 은닉층) 사이의 가중치 변화율은 식 2(혹은 식 3)와 같다.

$$E = \sum_k (d_k - y_k)^2 \quad \text{식 1}$$

$$w_{kh}(t+1) = w_{kh}(t) + \Delta w_{kh} + \alpha [w_{kh}(t) - w_{kh}(t-1)] \quad \text{식 2}$$

$$w_{hi}(t+1) = w_{hi}(t) + \Delta w_{hi} + \alpha [w_{hi}(t) - w_{hi}(t-1)] \quad \text{식 3}$$

2.3 대상유역

대상유역으로는 소양강댐유역으로 선정하였으며, 1011유역과 1012유역으로 구분하였다. 소양강댐유역중 1011유역의 면적은 931.22km², 1012유역의 면적은 1852.04km²이다. 그림 3은 대상유역의 위치와 대상유역 내의 AWS 관측소의 위치를 설명하고 있다.

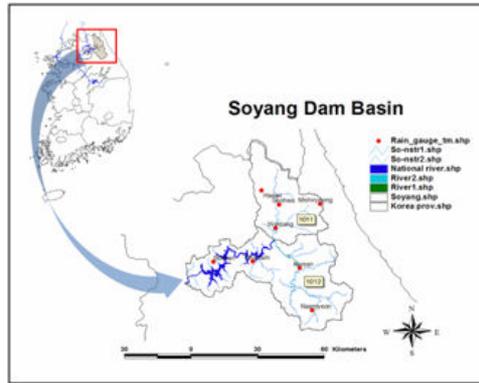
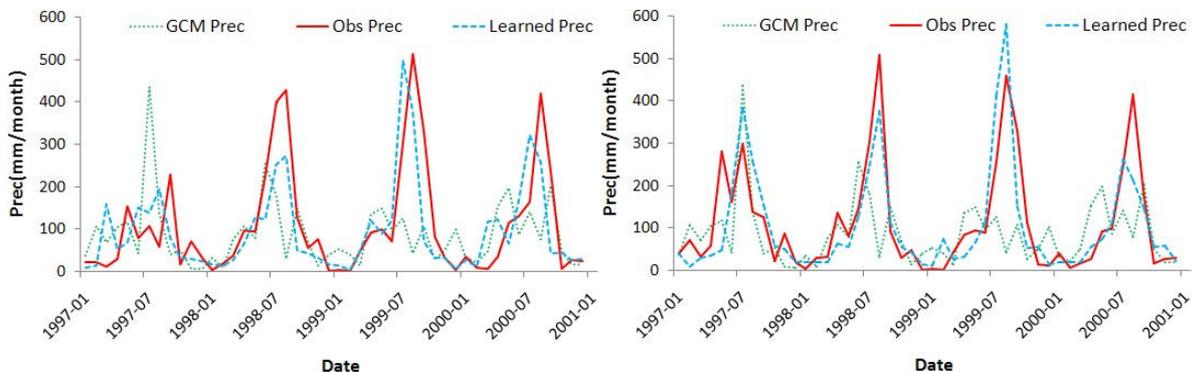


그림 2. 소양강댐 유역

3. 연구방법

대규모 대기모형의 모의결과를 이용하여 댐유역 스케일에서의 지역화 기법을 적요하기 위하여 1997년부터 2000년까지의 AWS의 관측값을 목표값으로 설정하고 역전파알고리즘을 이용하여 학습을 실시하였다. 인공신경망을 이용하여 스케일 상세화를 실시한 결과 GCM 강수에 비해서 향상된 모의값을 얻을 수 있었다. 1997년의 경우 GCM 강수와 관측강수의 연 총강수량의 차이는 314.22mm/yr였지만 인공신경망 학습 후 112.89mm/yr로 인공신경망을 이용하여 201.32mm/yr정도 향상된 모의값을 얻을 수 있었으며, 1999년의 경우 GCM 강수와 관측강수의 연 총강수량의 차이는 650.63mm/yr였지만 인공신경망 학습후 182.05mm/yr로 인공신경망을 이용하여 468.58mm/yr 향상된 모의값을 얻었다. 1997년의 경우 GCM 강수와 관측강수의 연 총강수량의 차이는 172.38mm/yr였지만 인공신경망 학습 후 64.78mm/yr로 인공신경망을 이용하여 107.59mm/yr정도 향상된 모의값을 얻을 수 있었으며, 1999년의 경우 GCM 강수와 관측강수의 연 총강수량의 차이는 551.38mm/yr였지만 인공신경망 학습후 97.65mm/yr로 인공신경망을 이용하여 453.73mm/yr 향상된 모의값을 얻었다.



(a) 월 누적강수의 학습성과(1011유역)

(b) 월 누적강수의 학습성과(1012유역)

그림 3. 월 누적강수의 학습성과(소양강댐 유역)

4. 연구결과

본 연구에서는 GCM이 고유적으로 가지고 있는 저해상도의 모의값을 유역규모 혹은 지역규모에서의 시공간적인 스케일특성을 부여함으로써 대상유역에서의 월 누적강수량과 일 평균온도를 예측하는데 있다.

각각의 대상유역에서 월 누적강수에 영향을 미치는 예측인자는 강수플럭스(GCM prec)가 6회로 모든 유역에서 월 누적강수를 학습하는데 가장 중요한 예측인자임을 확인할 수 있었다. 또한, 지표면으로부터 하강하는 대기 중의 장과플럭스(rlds)가 5회, 지표면으로부터 발생하는 현열(hfss)이 4회로 CGCM3.1/T63 20C3M에서 모의된 다른 기상변수에 비하여 발생빈도가 높게 나타났으며, 대규모 대기모형모의결과를 인공지능망의 입력값으로 활용한 결과 원시 GCM 결과값에 비하여 향상되었음을 확인할 수 있었다.

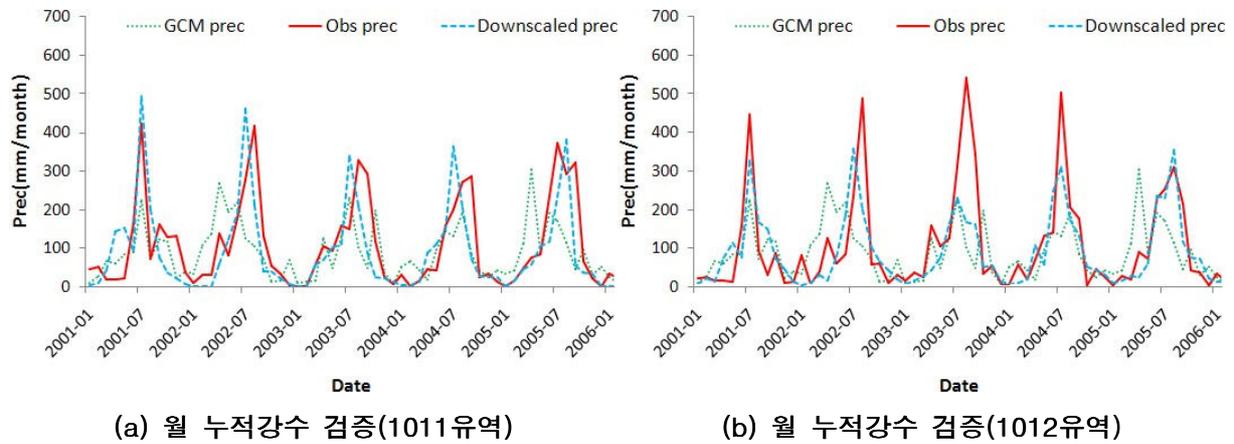


그림 4. 월 누적강수 검증(소양강댐 유역)

참 고 문 헌

- Dibike, Y.B., Coulibaly, P.(2006) Temporal neural networks for downscaling climate variability and extremes, *Neural Networks*, Vol.19, Issue 2, March 2006, Pages 135-144.
- Landman, W.A., Mason, S.J., Tyson, P.D. and Tennant, W.J.(2001) Statistical downscaling of GCM simulation to Streamflow, *Journal of Hydrology*, Vol. 252, issues 1-4, pp.221-236.
- Tripathi, S., Srinivas, V.V. and Nanjundiah, R.S.(2006) Downscaling of Precipitation for climate change scenarios: A support vector machine approach, *Journal of Hydrology*, Vol. 330, Issues 3-4, pp.621-640.
- Wilby, R.L., Hassan, H. and Hanaki, K.(1998) Statistical downscaling of hydrometeorological variables using general circulation model output, *Journal of Hydrology*, Vol. 205, issues 1-2, pp.1-19.
- Valverde Ramírez, M.C., de Campos Velho, H.F.(2005) Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *J. Hydrol.*, Vol. 301, issues 1-4, pp. 146-162.