

## 다층 퍼셉트론 인공신경망 모형을 이용한 가뭄예측 Drought Forecasting Using the Multi Layer Perceptron (MLP) Artificial Neural Network Model

이 주 현\* / 김 종 석\*\* / 장 호 원\*\*\* / 이 장 춘\*\*\*\*

Lee, Joo-Heon / Kim, Jong-Suk / Jang, Ho-Won / Lee, Jang-Choon

### Abstract

In order to minimize the damages caused by long-term drought, appropriate drought management plans of the basin should be established with the drought forecasting technology. Further, in order to build reasonable adaptive measurement for future drought, the duration and severity of drought must be predicted quantitatively in advance. Thus, this study, attempts to forecast drought in Korea by using an Artificial Neural Network Model, and drought index, which are the representative statistical approach most frequently used for hydrological time series forecasting. SPI (Standardized Precipitation Index) for major weather stations in Korea, estimated using observed historical precipitation, was used as input variables to the MLP (Multi Layer Perceptron) Neural Network model. Data set from 1976 to 2000 was selected as the training period for the parameter calibration and data from 2001 to 2010 was set as the validation period for the drought forecast. The optimal model for drought forecast determined by training process was applied to drought forecast using SPI (3), SPI (6) and SPI (12) over different forecasting lead time (1 to 6 months). Drought forecast with SPI (3) shows good result only in case of 1 month forecast lead time, SPI (6) shows good accordance with observed data for 1-3 months forecast lead time and SPI (12) shows relatively good results in case of up to 1~5 months forecast lead time. The analysis of this study shows that SPI (3) can be used for only 1-month short-term drought forecast. SPI (6) and SPI (12) have advantage over long-term drought forecast for 3~5 months lead time.

**Keywords** : artificial neural network, drought, multi layer perceptron, SPI, forecasting

### 요 지

장기간의 가뭄에 의한 피해를 최소화하기 위해서는 유역에 적합한 가뭄관리 대책의 수립과 함께 미래에 발생하게 될 가뭄을 미리 예측할 수 있는 기술이 구축되어야 한다. 또한 미래의 가뭄에 대한 합리적 대응 방안을 수립하기 위해서는 가뭄의 지속기간(duration)과 심도(severity)의 정량적인 예측이 선행되어야 한다. 본 연구에서는 수문 시계열의 예측에 가장 많이 이용되고 있는 대표적인 통계학적 기법인 인공신경망 모형(Artificial Neural Network Model)과 가뭄지수를 이용하여 남한지역의 서울, 대전, 대구, 광주 등의 4개 기상관측소를 선정하여 가뭄예측을 시도하였다. 가뭄 예측을 위하여

\* 교신저자, 중부대학교 공과대학 토목공학과 교수 (e-mail: leejh@joongbu.ac.kr, Tel: 041-750-6744)

Corresponding Author, Professor, Dept. of Civil Engineering, Joongbu University, Chungnam, Korea

\*\* 서울시립대학교 공과대학 토목공학과 연구교수 (e-mail: jongsuk@uos.ac.kr)

Research Prof., Dept. of Civil Engineering, University of Seoul, Seoul, Korea

\*\*\* 중부대학교 공과대학 토목공학과 석사과정 (e-mail: hs980216@hotmail.com)

Graduate Student, Dept. of Civil Eng., Joongbu University, Chung Nam, Korea

\*\*\*\* 전북대학교 공과대학 자원에너지공학과 교수 (e-mail: lj@jbnu.ac.kr)

Professor Dept. of Mineral Resources Energy Engineering Chonbuk National University, Jeonbuk, Korea

남한지역 내 선정된 기상관측소의 관측된 과거 강수량 자료를 이용하여 산정된 SPI (Standardized Precipitation Index)를 입력변수로 하여 다층 퍼셉트론(Multi Layer Perceptron) 인공신경망 모델에 적용하였으며, 매개변수 보정을 위한 학습기간으로 1976~2000년과 2001~2010년을 예측을 위한 검증기간으로 선정하여, 학습 및 예측을 시도하였다. 학습된 최적의 예측모형을 이용하여 서로 다른 선행예보시간(1~6개월)을 갖고 SPI (3), SPI (6), SPI (12)별로 가뭄을 예측하였으며, 가뭄예측 결과, SPI (3)의 경우에는 1개월 선행예보에서만 좋은 결과를 나타내었으며, SPI (6)의 경우 1~3개월 후의 가뭄을 예측하는 경우에 비교적 관측자료와 잘 일치하는 결과를 나타내었다. SPI (12)의 경우에는 약 5개월 후까지의 가뭄예측에 양호한 결과를 나타내었다.

**핵심용어** : 인공신경망, 가뭄, 다층 퍼셉트론, SPI, 예측

## 1. 서 론

가뭄이란 일정기간 동안의 무강우에 의해서 발생하는 기상현상이다. 가뭄 지속기간이 길어짐에 따라서 물부족이 발생하고, 이로 인하여 농업분야, 자연생태분야, 생공용수 관리 분야에 막대한 피해를 초래하게 되며 또한 가뭄은 홍수와 달리 상당한 시간을 갖고 발생하기 때문에 중, 장기적으로 가뭄을 예측한다는 것은 가뭄에 의한 재해에 대비하는 가장 첫 걸음이라고 할 수 있다. 즉, 가뭄이 발생하였을 때 가뭄의 지속기간 및 가뭄의 심도 등을 정확하게 예측할 수 있다면 가뭄에 의한 피해를 최소화 할 수 있을 것이다.

가뭄을 예측하기 위한 학술적, 기술적 시도가 많은 학자들에 의해서 시도 되었으며, 가뭄을 예측하고 전망하는 방법은 시나리오에 기반을 두는 방법과 실시간으로 가뭄을 예측하는 비시나리오 기반의 방법으로 구분될 수 있다.

시나리오에 기반을 두어 가뭄을 전망하는 방법은 기후변화 시나리오와 같은 미래의 기후 및 기상변화를 전망하고 이에 기반을 두어 중기 혹은 장기적으로 가뭄의 발생 빈도와 함께 가뭄의 시공간적인 분포 특성을 전망하는 방법으로 기후변화 시나리오 기반의 가뭄전망 연구의 경우 Kwon et al. (2009)은 3개월 GCM모델 예측 결과를 바탕으로 2009년도 PDSI 가뭄지수를 산정하여 가뭄심도에 대한 단기예측을 실시하였다. 그리고 Ghosh and Mujumdar (2007)은 SPI (12)를 이용하여 10년 단위별 미래 가뭄전망을 하였다. 또한 Kim et al. (2013)은 GCM모델에 의해 생산된 기상전망 자료와 SPI (Standardized Precipitation Index)를 이용한 빈도해석을 통해 SDF (Severity-Duration-frequency) 곡선을 유도한 후, 한반도의 확률가뭄우심도를 작성하여 미래 기후변화를 고려한 가뭄에 취약한 지역을 제시한 바 있다.

한편, 비시나리오 기반으로 가뭄을 전망하는 방법은 여러 가지의 통계학적 방법(statistical method)과 물리적 모

델(physical model)에 기반을 둔 확정론적 수치해석 방법을 이용하여 실시간으로 단기 혹은 중기의 미래 가뭄을 예측하는 것이다. 가뭄을 통계학적인 방법으로 예측하기 위해서 시도된 방법 중에서 가장 대표적인 방법이 다양한 수문시계열의 예측에 많이 이용되고 있는 ARIMA (Autoregressive Integrated Moving Average) 모형이다. 하지만 ARIMA 모형은 기본적으로 예측하고자 하는 시계열의 정상성 (Stationarity)을 전제로 하는 선형모형이라는 점에서 비정상성(non-stationary)과 비선형성(non-linearity)을 갖고 있는 수문시계열을 예측하는 데는 많은 한계를 갖고 있으며, 이분산성(Heteroskedasticity)일 경우 예측을 위해서는 관측치에 대해 가중치를 줌으로써 동일한 오차항 분산을 갖도록 해야 하는 단점이 있다. 또한 Belayneh and Adamowski (2012)는 SPI (3), (12)를 서포트 벡터 회귀기법과 웨이블릿 신경망에 적용하였으며, 예측구조는 RMSE (Root mean square error), MAE (Mean absolute error) 및 R (Correlation Coefficient)를 통해 선정하였고, 1~6개월의 선행예보 시간을 갖고 가뭄을 전망하였다. 그리고 Paulo et al. (2005)은 SPI를 이용하여, Markov chain 및 Loglinear 모델을 적용하여 가뭄예측의 정확도를 검증하였으며, Bacanlı et al. (2008)은 터키의 아나톨리아 지역을 대상으로 Neuro-Fuzzy 모델을 적용하여 1964~2006년 기간의 월별 평균 강수량과 SPI를 이용하여 가뭄을 예측하였다.

따라서 수문시계열의 비정상성과 비선형성을 고려한 예측을 시도하기 위해서는 비선형 모형의 적용이 검토되어야 하며 최근에 와서는 인공신경망 모형(Artificial Neural Network Model)을 이용하여 비선형 수문시계열을 예측하는 연구가 많이 시도되고 있다. 인공신경망(Artificial Neural Network)은 예측인자와 예측변수사이의 복잡한 비선형성을 고려하는데 유용하다. 최근에 와서는 다층 퍼셉트론(Multi Layer Perceptron, MLP) 인공신경망이 수문시계열의 예측에 자주 활용되고 있는데 예측을 하고자

하는 변수층의 개수가 증가할수록 퍼셉트론이 형성하는 결정 구역의 특성을 더욱 고급화 할 수 있는 네트워크가 바로 다층 퍼셉트론이다(Chen et al., 2012; Morid et al., 2007; Souhaib and Rob, 2012).

본 연구는 가뭄 예측에 대한 오차를 알고리즘과 결합하여 다층 퍼셉트론(MLP) 네트워크 구조를 인공신경망 모형에 적용하고, SPI를 입력 및 출력변수로 구성하여 가뭄 예측을 시도하였다. 역전파 알고리즘(Back Propagation Algorithm)을 통한 인공신경망의 학습을 통해 오차를 최소화하고 가뭄예측에 적합한 신경망 구조를 선별하여 가뭄예측을 시도함으로써, 서로 다른 강수의 지속기간으로부터 산정된 SPI와 인공신경망의 가뭄예측에 대한 적용성을 평가하고자 하였다.

## 2. 자료의 구축

### 2.1 기상 관측소 및 자료 기간

본 연구에서는 가뭄을 예측하기 위해 기상청의 기상관측소 중에서 자료의 보유기간이 비교적 길고 한강, 낙동강, 금강, 영산강유역에 위치하는 서울, 대전, 대구, 광주 기상관측소를 선정하였다. 선정된 4개 기상관측소의 과거 관측된 강수량자료(1976~2010년)를 이용하여 가뭄지수를 산정하였다.

### 2.2 가뭄지수

가뭄을 정량적으로 해석하기 위해서 가뭄지수를 활용하는 방법이 많이 이용되고 있으며, 본 연구에서는 SPI를 가뭄판단을 위한 가뭄지수로 선정하였다. 가뭄은 물의 수요에 비해 상대적으로 물의 부족을 유발하는 강수량의 감소에 의해 시작된다는 것에 착안하여 SPI가 개발되었다. 즉, 강수량이 부족하면 용수 공급원인 지하수, 저수지 저류량, 토양함유수분, 하천유출량 등에 각기 다른 영향을 미친다는 가정으로부터 개발한 것이다. SPI는 특정한 시간에 대한 계산 시간 누적단위를 1, 3, 6, 12개월 등과 같이 설정하고, 시간단위별로 강수부족량을 산정하여 가뭄지수를 계산함으로써 개개의 용수공급원이 가뭄에 미치는 영향을 산정하는 효율적인 방법이다(Mckee et al., 1993, 1995).

Fig. 1은 서울, 대전, 대구, 광주 기상관측소의 강수량자료(1976~2010년)를 이용하여 서로 다른 지속기간을 갖는 SPI (3), SPI (6) 및 SPI (12)를 산정하여 나타낸 것이고, Table 1은 SPI에 의한 가뭄의 분류 기준을 나타내고 있다.

Table 1. SPI Drought Categories (Mckee et al., 1993)

SPI Range	Moisture Condition
2.00 ≤	Extremely Wet
1.50 ~ 1.99	Very Wet
1.00 ~ 1.49	Moderately Wet
-0.99 ~ 0.99	Near Normal
-1.00 ~ -1.49	Moderately Dry
-1.50 ~ -1.99	Severely Dry
≤ -2.00	Extremely Dry

Fig. 1에서 나타난 것과 같이, SPI의 지속기간이 짧은 SPI (3)의 경우에는 습윤(Wet)과 건조(Dry) 상태를 오가는 주기가 매우 짧지만 반대의 경우인 SPI (12)의 경우에는 12개월의 누적강수를 이용하여 지수를 산정하기 때문에 습윤과 건조를 오가는 주기가 매우 긴 것으로 분석되었다.

그러나 3가지 지속기간의 SPI는 지역별로 약간의 차이는 있지만 모두 과거의 대표적인 가뭄년도(1977~1978, 1988, 1994~1995, 2001, 2008~2009)를 잘 재현하고 있는 것으로 나타났다.

### 2.3 가뭄지수의 정상성 판단

수문시계열의 비정상성과 비선형성을 고려한 예측을 위해서 본 연구에서는 ANN을 이용하고자 하며, 예측을 위한 입력데이터로 가뭄지수를 활용하기 전에 과거 관측된 강수량을 토대로 산정된 가뭄지수 시계열의 정상성을 판단하고자 한다.

정상성이란 시계열 자료의 수준(level) 또는 평균이 일정하고, 자료 사이에 변화하는 폭에 증감이 없이 일정하며, 시간과 무관하게 일정한 간격을 두고 떨어져 있는 두 변수들의 공분산이 동일하다는 3가지 조건을 만족할 때 이를 정상성이라 한다. 본 연구에서는 가뭄지수의 정상성을 판단하기 위하여 시계열 자료의 경향성 또는 추세 분석방법 중 Mann-Kendall 검정 방법을 실시하였고(Jung et al., 2007), 자기상관계수(Auto Correlation Function)에 의한 correlogram을 통하여 가뭄지수의 정상성을 판단하고자 한다.

#### 2.3.1 Mann-Kendall 검정

Mann-Kendall Test는 시계열의 추세를 분석하기 위한 대표적인 비모수추정 기법으로서, 다른 통계학적 검정 방법들이 표본크기, 분산, 왜도와 같은 표본자료특성에 영향

을 받는데 반해 표본자료의 특성에 민감하지 않고 간단하면서도 결측치나 감지 한계를 벗어나는 자료에 대해서도 적용이 가능하다(Seo, 2011).

Table 2는 Eq. (1)을 통해 검정한 가뭄지수별(1976~2010) 4개 관측소에 대한 Mann-Kendall Test 결과를 정리하였다. Table 2에서 알 수 있듯이, 서울관측소의 경우

는 SPI (6)과 (12)에서 상당한 유의성을 갖고 증가추세인 것을 알 수 있으며 대전과 광주 관측소는 유의한 수준은 아니지만 시계열 자료가 증가하는 추세를 갖고 있음을 알 수 있다. 또한, 대구관측소의 경우에는 SPI (3)과 (12)에서 감소하는 경향을 보였고 SPI (6)에서는 증가하는 경향을 나타내었지만 유의할 만한 수준은 아닌 것으로 나타났다.

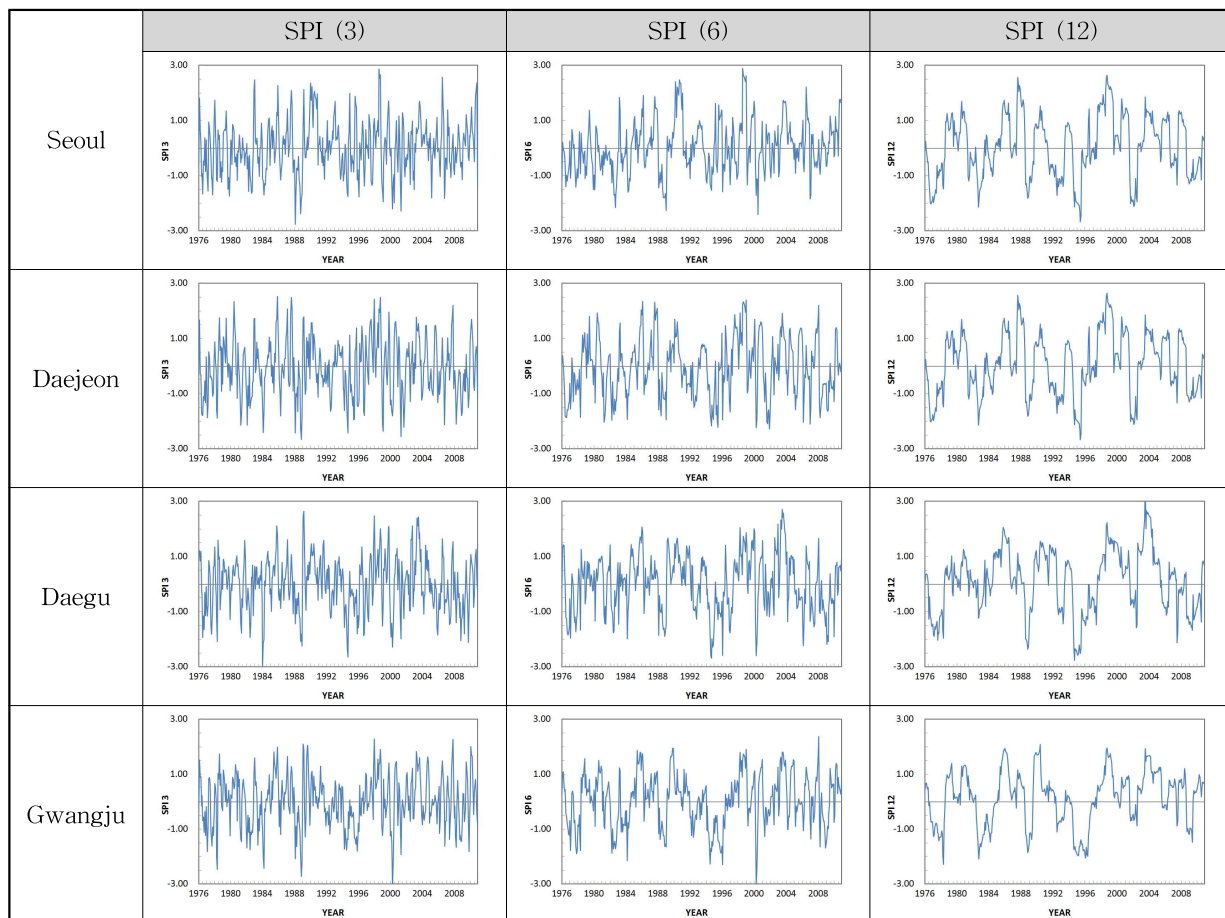


Fig. 1. SPI Time Series (1976-2010) for 4 Different Weather Stations

Table 2. Trend Analysis for SPI Time Series (1976-2010) using Mann-Kendall Test

	SPI (3) Confidence limit:90%		SPI (6) Confidence limit:90%		SPI (12) Confidence limit:90%	
	MK-Statistics	Trend	MK-Statistics	Trend	MK-Statistics	Trend
	Seoul	1.48	△	2.30	↑	2.91
Daejeon	0.44	△	0.70	△	1.09	△
Daegu	-0.44	▼	0.54	△	-0.28	▼
Gwangju	0.17	△	0.74	△	1.45	△

↑: Increasing trend that satisfy the confidence level

↓: Decreasing trend that satisfy the confidence level

△: Increasing trend that does not satisfy the confidence level

▼: Decreasing trend that does not satisfy the confidence level

○: No trend

$$T = \sum_{j < i} \text{sign}(x_i - x_j) \quad 1 \leq j < i \leq n \quad (1)$$

여기서,  $T$ =시계열에 대한 일변량 Mann-Kendall 통계치

$$\text{sign}(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x = 0 \\ -1, & \text{if } x < 0 \end{cases}$$

### 2.3.2 상관도 분석(Correlogram Test)

Mann-Kendall 검정을 통한 경향성 분석 외에도 수문 시계열의 지속성과 정상성 여부를 판단하기 위한 방법으로 상관도(correlogram)를 이용하는 방법이 있다. 수문 시계열의 지체량에 대한 자기상관계수를 Eq. (2)에 의해서 산정한 뒤, 지체량에 대해서 도시한 것이 Fig. 2이다. 정상시계열의 경우에는 지체량이 증가함에 따라서 자기상관계수가 급격히 0으로 수렴하지만, Fig. 2에서 알 수 있듯이 SPI (6)에 대한 상관도는 뚜렷한 주기성을 보임과 동시에 신뢰한계를 벗어나고 있는 비정상시계열임을 알 수 있다.

$$\gamma_k = \frac{C_k}{C_0} = \frac{\sum_{t=1}^{N-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^N (X_t - \bar{X})^2} \quad (2)$$

여기서,  $\gamma_k$ =지체량  $k$ 에 대한 변량간의 자기상관계수(Auto-correlation coefficient)

$C_k$ =지체량이  $k$ 일 때의 자기공분산(Autocovariance)

$C_0$ =지체량이 0일 때의 자기공분산(Autocovariance)

$X_t$ =주기가  $t$ 일 때의 원시계열  $X$ 년과 주기의 평균

## 3. 가뭄전망을 위한 인공신경망 구축

### 3.1 가뭄전망 프로세스

Fig. 3은 인공신경망을 이용하여 가뭄을 예측하기 위한 절차를 순서대로 나타내고 있으며, 가뭄예측을 위해서는 크게 데이터 구축, 학습 및 매개변수 조정, 검증 및 예측 등 크게 3가지 단계로 분류된다. 본 연구에서는 과거에 관측된 강수량자료를 이용하여 산정된 SPI를 학습기간(1976~

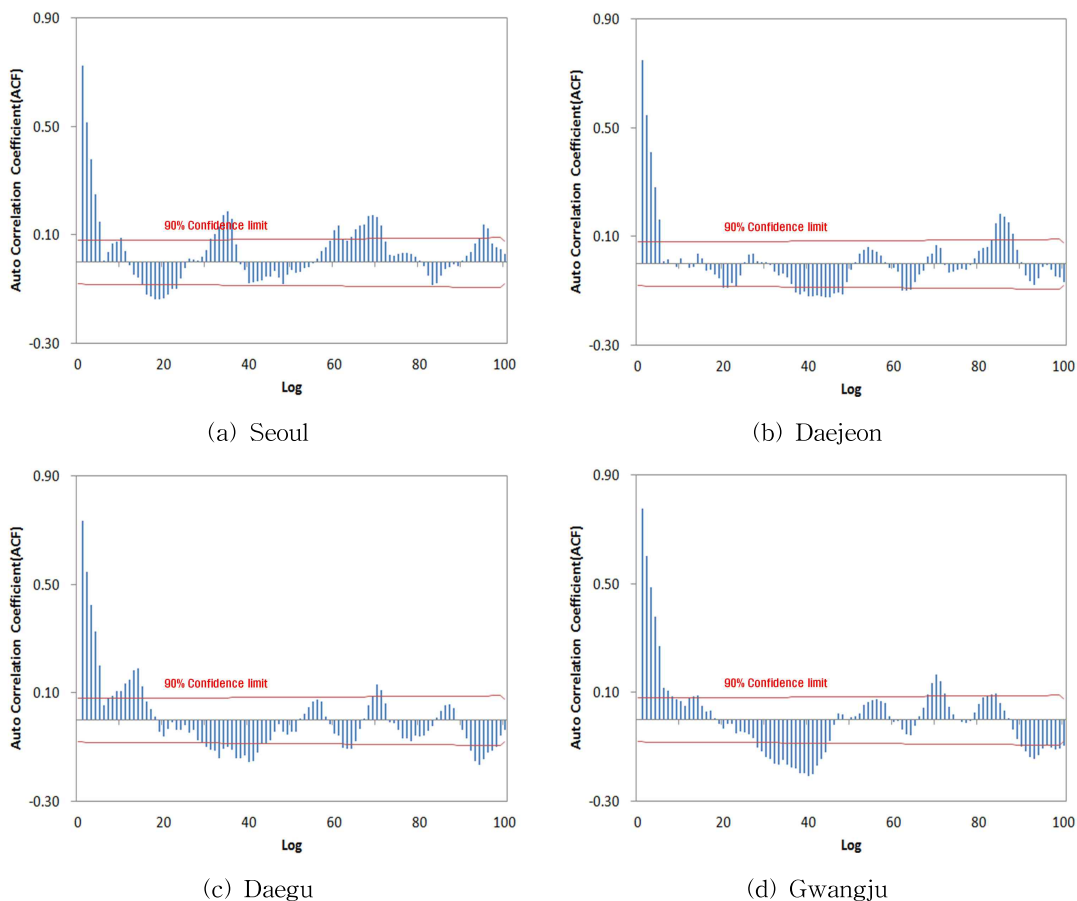


Fig. 2. Correlograms for SPI (6) Time Series

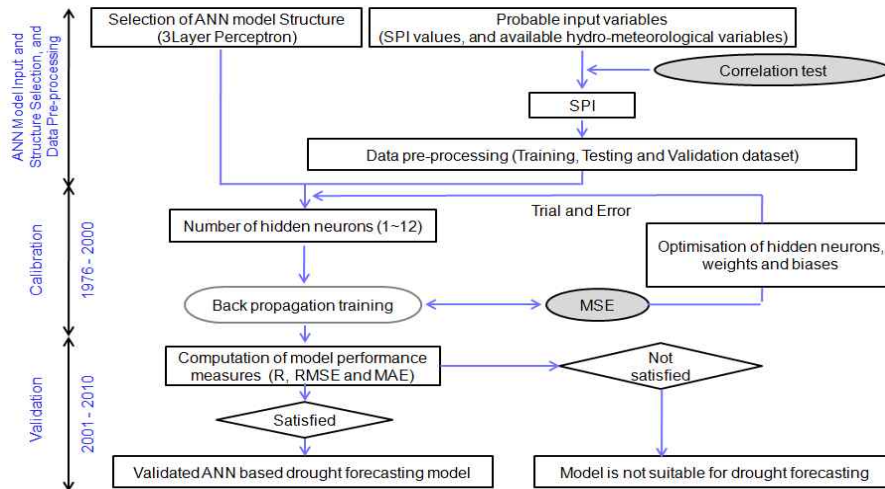


Fig. 3. Flowchart of ANN based Drought Forecasting Model Development Process

2000)과 예측을 위한 검증기간(2001~2010)으로 분류하였다.

학습단계에서는 역전파 알고리즘(Back Propagation Algorithm)을 이용하여 입력변수와 출력변수간의 비선형적 관계를 학습함으로써 신경망의 구조를 결정하며, 검증단계에서는 학습을 통해 결정된 구조로 예측을 시도하고 Eqs. (3) and (4)에 제시한 RMSE(Root mean square error), MAE(Mean absolute error) 및 R(Correlation Coefficient) 등에 의한 모형의 성능오차를 분석하여 가뭄예측의 정확도를 검증한다.

$$RMSE = \sqrt{\frac{1}{p} \sum_{i=1}^p [(X_o)_i - (X_f)_i]^2} \quad (3)$$

$$MAE = \frac{1}{p} \sum_{i=1}^p |(X_o)_i - (X_f)_i| \quad (4)$$

여기서,  $p$ : SPI의 총 개수,  $X_o$ : 관측된 SPI 표준화 값,  $X_f$ : 예측된 SPI 표준화 값

### 3.2 다층 퍼셉트론 인공신경망(Multi Layer Perceptron Neural Network)

인공신경망은 시냅스의 결합으로 네트워크를 형성한 인공 뉴런(노드)이 학습을 통해 시냅스의 결합 세기를 변화시켜, 문제 해결 능력을 가지는 모델 전반을 가리킨다.(Kang and Lee, 2011). 인공신경망 모형에는 여러 구조가 있으며, 그 중 퍼셉트론은 입력층, 출력층, 가중치 및 임계치 벡터로 구성되어있다.

퍼셉트론의 특징으로는 사용되는 은닉층의 활성화함수가 출력값의 경우 -1과 1사이 값만을 반환되기 때문에 기상학적 변수에 의한 비선형성을 고려해야 하는 SPI의 경

우 은닉층에 활성화 함수에 적용하기 위해서는 데이터를 가공할 필요가 있다. 그러므로 SPI를 Eq. (5)와 같이 SPI를 0과 1사이 값에 갖도록 데이터를 변환하여 입력 변수로 사용하였다.

$$X_o = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (5)$$

여기서,  $X_o$ : 표준화를 통해 정규화 된 값,  $X$ : 관측된 SPI 값,  $X_{\min}$ : 관측된 SPI 최소값,  $X_{\max}$ : 관측된 SPI 최대값  
이렇게 표준화 된 입력변수는 Eq. (6)과 같이 가중치와 곱하고 임계값을 더하여 입력값을 계산된다.

$$I_j = \sum w_{ji} X_o + B_j \quad (6)$$

여기서,  $I_j$ : 입력값,  $w_{ji}$ : 가중치 벡터,  $X_o$ : 입력층의 데이터,  $B_j$ : 임계값 벡터

계산된 입력값은 활성화 함수로 보내지며, Eq. (7)과 같이 은닉층의 입력데이터를 계산하게 되며, 신경망에서의 활성화 함수는 입력층 데이터에 따라 여러 함수가 존재한다. 신경망에서 사용되는 대표적인 활성화 함수는 시그모이드(sigmoid)형, 계단함수(hard limiter), 임계논리(threshold logic)함수 등이 있으며(Jeong et al., 2002), 본 연구에서는 시그모이드 함수를 사용하였다.

$$y_d = f(I_d) \quad (7)$$

여기서,  $y_d$ : 은닉층의 입력데이터,  $f$ : 활성화 함수

은닉층의 경우 활성화 함수에서 계산된 은닉층의 입력 데이터와 가중치를 곱하고 임계값을 더하여 Eq. (8)과 같이 계산된다. 계산된 값은 출력층으로 보내지게 된다.

$$L_k = \sum w_{kj}y_d + B_k \quad (8)$$

여기서,  $L_k$ : 반응벡터,  $w_{kj}$ : 은닉층의 가중치 벡터,  $B_k$ : 은닉층에서의 임계값

이처럼 다층 퍼셉트론은 단층 퍼셉트론과 유사한 구조를 가지고 있지만 여러 개의 층으로 구성되어 있으며, 역전파 알고리즘을 통해 단층 퍼셉트론에서 하나의 조정층으로 구성되는 모델들의 한계점들과 선형 분리 문제를 해결하였다.

역전파 알고리즘이란 인공지능망 학습방법으로 일반적으로 많이 이용되고 있으며, 학습을 하기 위해서는 입력 데이터와 원하는 출력 데이터가 존재해야 한다.

학습은 Eqs. (6)~(8)과 같이 단계별로 계산하게 되며, 마지막으로 Eqs. (3) and (4)와 같이 출력된 값과 관측된 값을 비교하여 관측된 값과의 차이를 줄이기 위해 반복하여 계산하는 과정을 역전파 알고리즘이라 한다.

본 연구에서는 단층 퍼셉트론의 단점을 보완한 다층 퍼셉트론을 선정하였으며, 비선형적인 예측에 적합한 역전파 알고리즘 방법을 이용하여 가뭄을 예측의 적용성을 평가하고자 한다.

### 3.3 네트워크 디자인 및 가뭄예측

다층퍼셉트론 인공지능망 모형을 이용하여 가뭄을 예측하기 위한 네트워크 구축을 위한 방법으로 DMSNN(Direct Multi-Step Neural Network) 방법이 있으며, DMSNN은 ARIMA 모델에 비해 신경망의 장점을 기반으로 하고 있다(Mishar and Desai, 2006).

DMSNN은 Fig. 4와 같이 여러 개의 입력층과 은닉층 및 출력층을 구성하고 있으며, 한 번에 여러 개의 출력층을 통해 각각의 출력값을 구하는 방식이다. 본 연구에서는 DMSNN 방법을 이용하여, 1~12개월까지의 가뭄을 예측하였다.

## 4. 결과의 분석

### 4.1 학습결과

본 연구에서는 다층 퍼셉트론 인공지능망에서 DMSNN

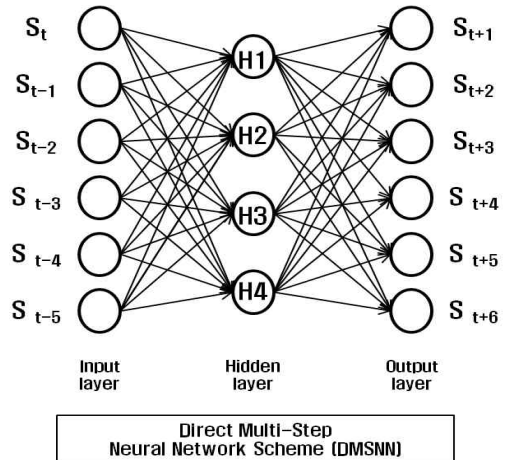


Fig. 4. Architecture of Three-Layer Perceptron Direct Multi-Step Neural Network (DMSNN)

방법으로 학습횟수를 100~10,000까지 설정하였으며, 입력층에 변수 또한 6개월에서 12개월까지 증가하여 학습을 실시하였다.

학습을 통한 결과는 일반화 판단을 통해 선택오차(select error)와 검정오차(test error)의 차이가 적은 구조를 선별하였으며, 선별된 구조 중 RMSE, MAE, R값을 통한 최적의 구조를 선정하여, 가뭄을 예측하였다. 예측결과 학습 횟수를 1,000회로 설정하고 입력층에 변수를 12개월로 입력하였을 때 학습결과가 가장 좋은 것을 나타냈다. Table 3은 여러 학습을 통해서 적합하다고 판단된 12개월 입력 변수와 학습 횟수 1,000번으로 학습을 했을 때에 서울 관측소에 SPI (3), SPI (6), SPI (12)에 대한 구조를 나타내고 있다.

### 4.2 가뭄예측 결과

본 연구에서는 다층퍼셉트론 인공지능망 모형을 적용하여 서울, 대전, 대구, 광주 관측소에 대하여 SPI를 이용한 가뭄을 예측하였다. 가뭄예측은 1976~2000년까지의 자료를 이용하여 학습된 구조를 기반으로 하여 2001~2010년 기간 동안에 대한 SPI를 예측하는 것으로 하였다. 4개 관측소 중 서울 관측소에 대한 결과를 Figs. 4~6에 나타내었으며, Table 3은 4개 관측소의 예측결과에 대한 오차 및 상관계수를 각 관측소 별로 정리하였다.

Table 3. Back Propagation Training Results and Model Summary

	SPI	Model Profile	Training	Number of Input Layer	Number of Hidden Layer
Seoul	3	MLP12:12-13-12:12	Back Propagation	12	13
	6	MLP12:12-16-12:12	Back Propagation	12	16
	12	MLP12:12-16-12:12	Back Propagation	12	16



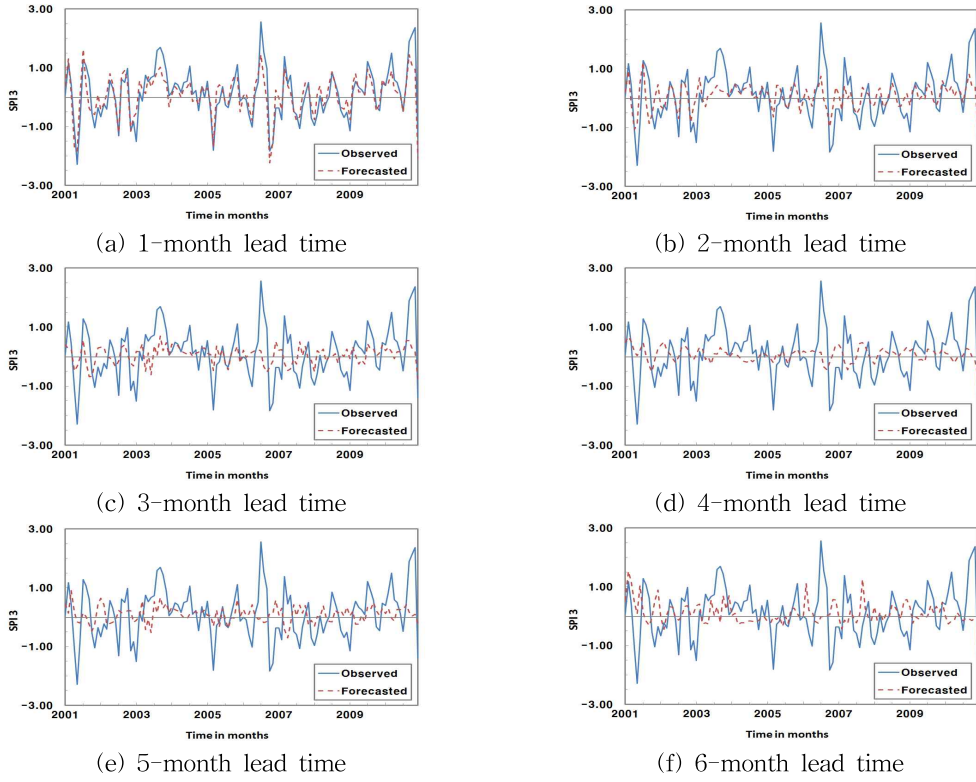


Fig. 5. Comparisons between Forecasted and Observed SPI (3) (2001~2010) over Different Forecasting Lead Time

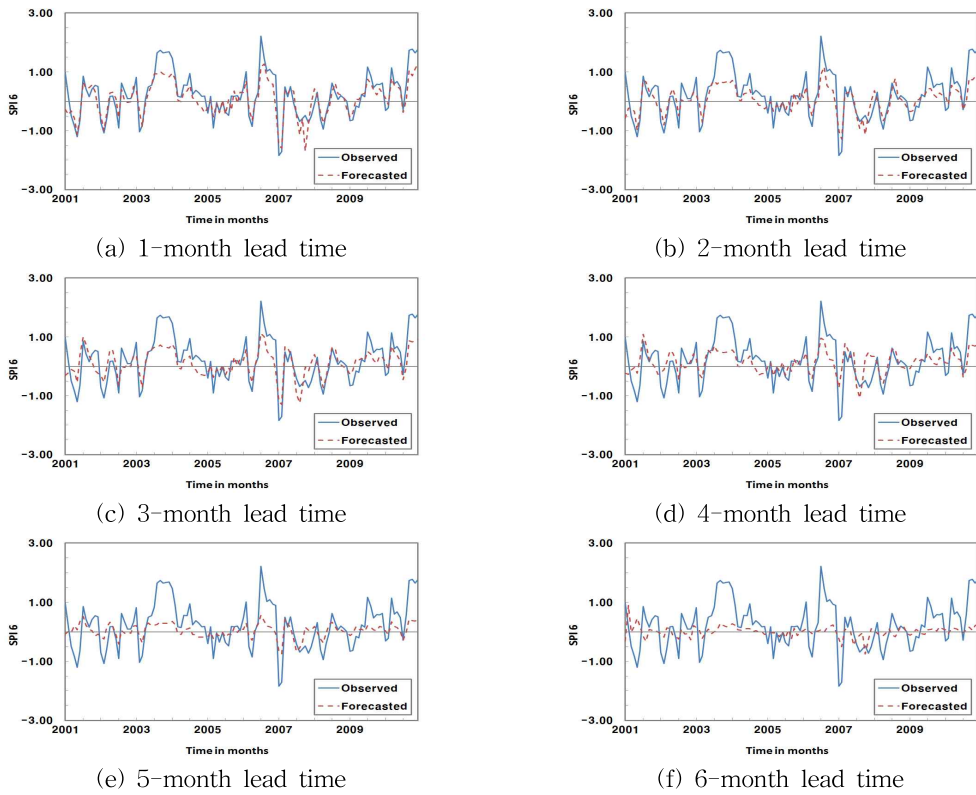


Fig. 6. Comparisons between Forecasted and Observed SPI (6) (2001~2010) over Different Forecasting Lead Time



#### 4.2.1 SPI 지속기간별(3, 6, 12개월) 가뭄 예측 결과

Fig. 5는 SPI (3)을 이용하여 서울관측소에 대하여 1~6개월까지 선행예보 시간을 갖고 가뭄을 예측한 결과이다. 예측한 결과를 보면, 1개월 선행예보의 경우에 예측치가 관측치와 매우 잘 일치하는 것으로 나타났다. 그러나 2개월 이상의 선행예보에서는 예측치의 정확도가 현격하게 감소하는 것으로 나타났다. Fig. 6은 서울 관측소에 대하여 1~6개월까지 선행예보 시간을 갖고 SPI (6)를 이용하여 가뭄 예측한 결과이며, SPI (6)의 경우 1~2개월까지 선행예보에 대한 예측치와 관측치가 일치하는 것으로 나타났다. 그러나 2개월 후의 선행예보에는 SPI (3)과 결과와 마찬가지로 급격하게 떨어지는 것으로 나타났으며, Fig. 7은 SPI (12)를 이용하여 서울 관측소에 대하여 가뭄 예측한 결과이다. SPI (12)의 경우 1~5개월의 예측치가 관측치와 일치하는 것으로 나타났으며, 이후의 선행예보에서는 예측치의 정확도가 다른 지수들과 마찬가지로 감소하는 것으로 나타났다. Table 4는 서울 관측소에 대한 지수별 오차와 상관성을 수치로 나타내고 있다. SPI (3), (6), (12)모두 1개월 선행예보 이후로 RMSE, MAE가 증

가하고, R이 감소하는 것을 볼 수 있으며, 이를 통해 선행예보 기간이 길어질수록 예측오차가 크며 상관성이 낮아지는 것으로 나타났다.

#### 4.2.2 가뭄등급별 예측결과 분석

Fig. 8은 4개의 기상관측소(서울, 대전, 대구, 광주)의 선행예보 시간(1~6개월)에 따른 SPI (3)에 대한 모의결과를 관측치와 비교하여 분석한 결과를 보여주고 있다. 4개 관측소 모두 선행예보시간이 길어질수록 SPI (3)에 대한 예측오차가 증가하는 것으로 나타나고 있다. 예측 검증기간(2001~2010) 전체에서 1개월 선행예보의 예측치가 관측치와 가장 잘 일치하는 것으로 나타났다. Fig. 8에서 보는바와 같이, 가뭄예측결과는 선행 예보기간에 따라 약간의 차이를 보이고 있으며 가뭄등급별로 상대적인 오차가 큰 것으로 나타났다. 본 연구에서는 관측치 가뭄자료의 분포특성을 고려하여 가뭄등급을 분류하였다. 중장기 가뭄을 대표하는 SPI (6)과 SPI (12)의 경우, 가뭄의 심도가 평활화되어 가뭄상태(Dry)인  $SPI \leq -1$ 의 발생이 상대적으로 적게 나타나 통계적 특성을 평가하기에는 무리가 있다. 따라서 단기가뭄을 나타내는 SPI3을 대상으로 3개의 가뭄등급(Dry:  $SPI3 \leq -1$ , Normal:  $-1 < SPI3 < 1$ ,

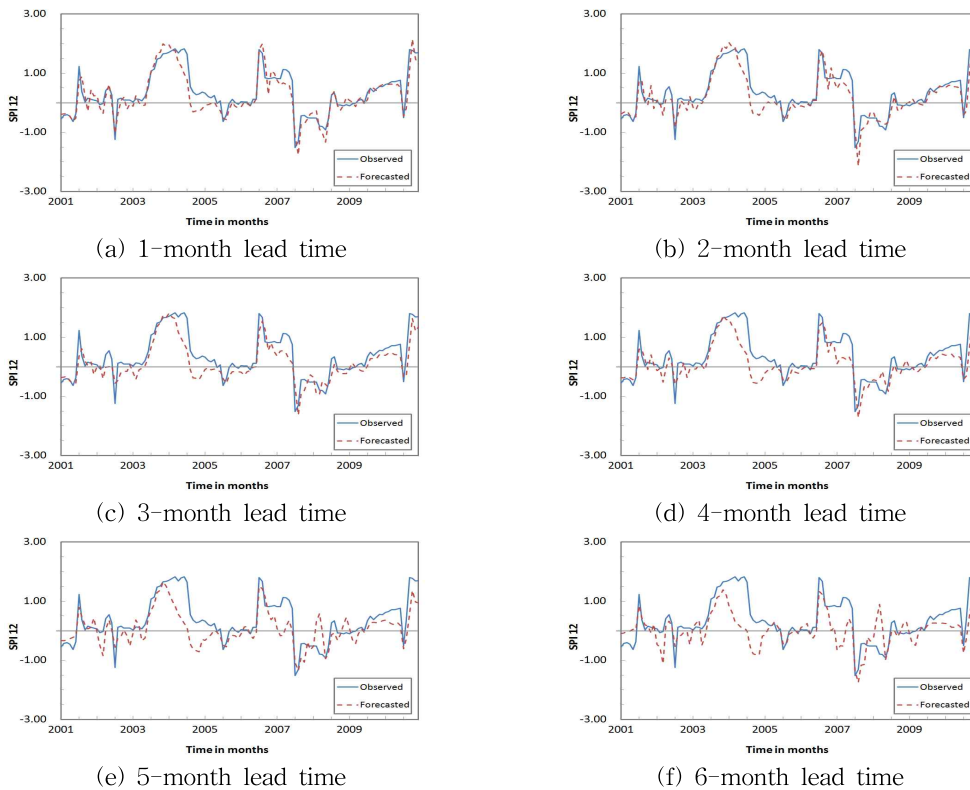


Fig. 7. Comparisons between Forecasted and Observed SPI (12) (2001~2010) over Different Forecasting Lead Time

Table 4. Forecasting Measures between Observed and Predicted SPI (2001~2010)

	SPI	Forecasting Measure \ Lead Time	1 month	2 month	3 month	4 month	5 month	6 month
Seoul	SPI (3)	RMSE	0.73	0.85	0.91	0.91	0.86	0.86
		MAE	0.45	0.54	0.55	0.54	0.52	0.58
		R	0.58	0.26	0.04	-0.08	0.10	0.15
	SPI (6)	RMSE	0.62	0.77	0.82	0.81	0.79	0.78
		MAE	0.47	0.60	0.64	0.62	0.61	0.63
		R	0.62	0.32	0.21	0.15	0.16	0.22
	SPI (12)	RMSE	0.52	0.63	0.67	0.69	0.71	0.74
		MAE	0.35	0.43	0.49	0.52	0.55	0.59
		R	0.77	0.65	0.61	0.57	0.56	0.52
Daejeon	SPI (3)	RMSE	0.84	0.94	1.00	1.00	1.00	1.00
		MAE	0.69	0.75	0.80	0.81	0.80	0.77
		R	0.53	0.28	0.01	-0.06	-0.08	-0.08
	SPI (6)	RMSE	0.74	0.93	0.99	1.11	1.13	1.10
		MAE	0.58	0.74	0.81	0.91	0.93	0.96
		R	0.72	0.50	0.38	0.17	0.05	-0.03
	SPI (12)	RMSE	0.55	0.75	0.76	0.79	0.86	0.93
		MAE	0.39	0.52	0.52	0.57	0.65	0.91
		R	0.85	0.69	0.67	0.63	0.53	0.44
Daegu	SPI (3)	RMSE	0.81	1.00	1.02	1.01	1.04	1.05
		MAE	0.63	0.81	0.84	0.83	0.86	0.83
		R	0.61	0.28	0.13	0.11	-0.04	-0.03
	SPI (6)	RMSE	0.88	1.01	1.07	1.15	1.20	1.22
		MAE	0.63	0.83	0.89	0.95	0.99	1.04
		R	0.69	0.48	0.43	0.35	0.22	0.14
	SPI (12)	RMSE	0.58	0.72	0.82	0.87	0.96	1.04
		MAE	0.35	0.52	0.60	0.66	0.78	1.16
		R	0.88	0.80	0.74	0.71	0.64	0.60
Gwangju	SPI (3)	RMSE	0.79	1.01	1.05	0.97	0.97	0.99
		MAE	0.63	0.81	0.83	0.77	0.79	0.75
		R	0.60	0.25	0.16	0.21	0.18	0.11
	SPI (6)	RMSE	0.70	0.92	1.04	1.06	1.15	1.05
		MAE	0.51	0.70	0.83	0.85	0.95	0.87
		R	0.67	0.39	0.26	0.18	0.02	-0.06
	SPI (12)	RMSE	0.40	0.53	0.57	0.63	0.68	0.72
		MAE	0.27	0.35	0.42	0.48	0.52	0.63
		R	0.86	0.75	0.71	0.63	0.55	0.49

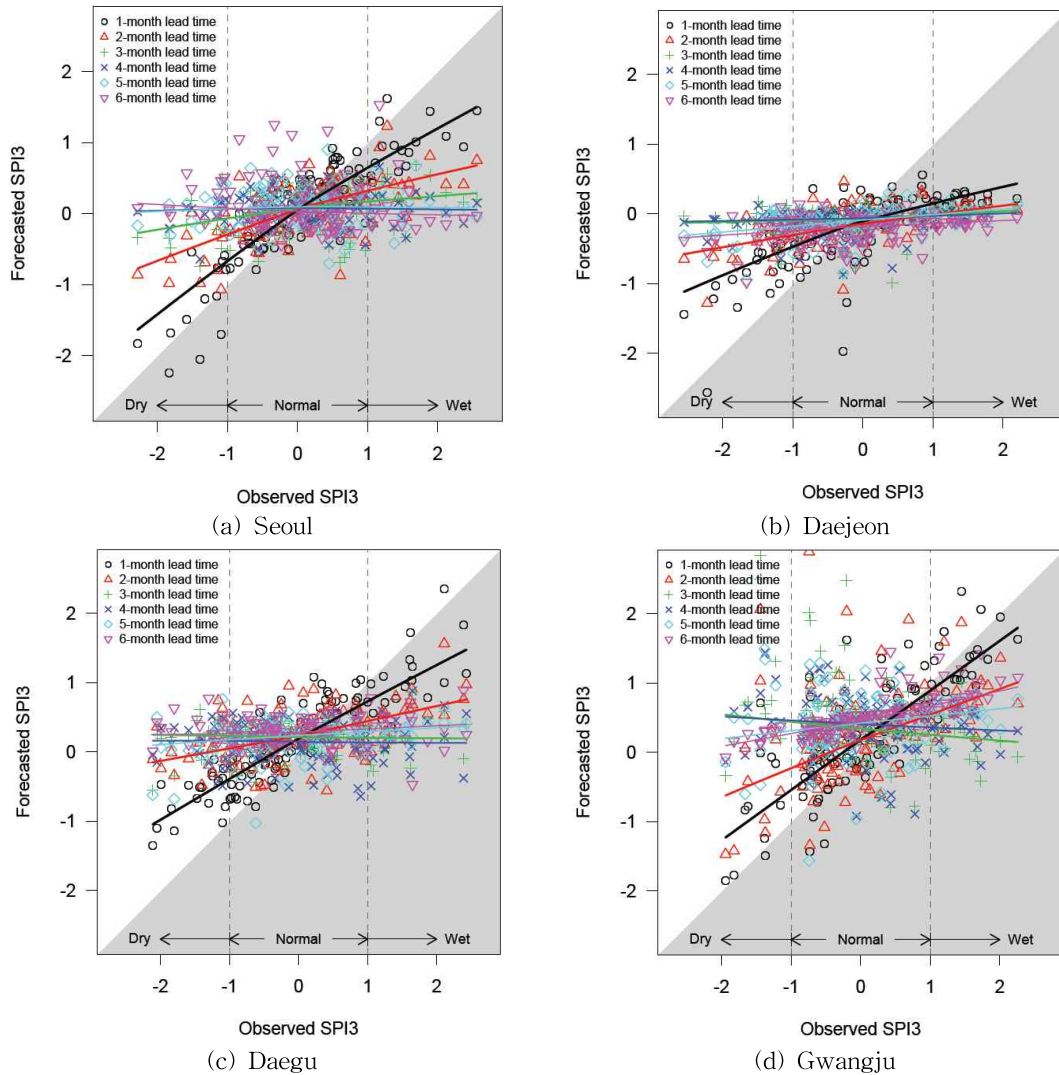


Fig. 8. Comparisons between Observed SPI3 and Forecasted SPI3 according to Different Conditions of Lead Time (2001~2010)

Wet:  $SPI3 \geq 1$ )으로 분류하였으며, 선행예보 기간과 가뭄등급에 따른 MLP ANN모형의 가뭄예측의 적용성을 평가하였다.

가뭄등급별 관측치와 예측치의 통계지표를 활용한 비교 결과는 Table 5에서 확인할 수 있다. 4개 관측소의 과거 가뭄(2001~2010)기간의 SPI (3)은 -2.55에서 2.56 사이에 분포하고 있으며, 전체 검증기간(120개월) 중 65.0~75.8%에서 Normal ( $-1 < SPI3 < 1$ ) 상태가 나타났고, Wet ( $SPI3 \geq 1$ , 13.3~17.5%) 및 Dry ( $SPI3 \leq -1$ , 9.2~18.3%) 상태는 상대적으로 적게 발생되었다. 선행예보 기간에 따라 차이는 있지만 전체적으로 대상 관측소에서 Dry조건( $SPI3 \leq -1$ )에서는 MLP ANN모형의 예측결과가 관측치와 비교하여 과대 추정되는 경향을 보이는 것으로 분석되었다.

그러나 Wet조건( $SPI3 \geq 1$ )에서는 Dry조건( $SPI3 \leq -1$ )과는 상반된 결과를 보이는 것으로 나타났다. 본 연구에서 적용한 MLP ANN모형의 가뭄예측 결과는 Normal상태를 잘 모의하는 것으로 확인되었으며, Wet 조건보다는 Dry조건을 더 잘 예측하는 것으로 나타났다. 또한 선행예보 시간이 짧을수록 가뭄예측에 대한 신뢰도가 높은 것으로 분석되어 장기예측보다는 단기 가뭄 예측에 활용하는 것이 좋을 것으로 판단된다.

#### 4.2.3 홍수기와 갈수기의 예측결과 비교 분석

Table 6은 서울관측소에 대하여 SPI가 예측된 결과를 홍수기(6, 7, 8, 9월)와 갈수기(1, 2, 3, 4, 5, 10, 11, 12월)로 분류하여 각 지수별로 비교하여 나타내고 있다. 예측결과를 분석

**Table 5. Statistical Comparison between Observed SPI3 and Forecasted SPI3 Based on the Moisture Conditions (2001~2010)**

Moisture Condition	Station	Statics	Lead Time (month)						Number of Events
			1	2	3	4	5	6	
Dry (SPI3 ≤ -1)	Seoul	RMSE	0.46	0.95	1.28	1.45	1.49	1.63	13
		MAE	0.38	0.85	1.22	1.40	1.46	1.59	
		R	0.62	0.35	0.36	0.33	0.49	0.25	
	Daejeon	RMSE	0.77	1.11	1.46	1.43	1.31	1.25	22
		MAE	0.72	1.05	1.39	1.37	1.25	1.19	
		R	0.74	0.55	0.03	0.16	0.47	0.44	
	Daegu	RMSE	0.88	1.55	1.69	1.60	1.58	1.81	18
		MAE	0.82	1.50	1.65	1.53	1.54	1.77	
		R	0.59	0.33	0.47	-0.16	0.46	-0.06	
	Gwangju	RMSE	0.99	1.60	2.55	2.15	1.96	1.70	11
		MAE	0.76	1.27	2.40	2.11	1.86	1.69	
		R	0.55	0.37	0.34	0.49	0.38	0.82	
Normal (-1 < SPI3 < 1)	Seoul	RMSE	0.34	0.47	0.55	0.55	0.57	0.62	91
		MAE	0.27	0.36	0.44	0.47	0.47	0.50	
		R	0.73	0.37	0.06	-0.09	-0.04	-0.05	
	Daejeon	RMSE	0.48	0.49	0.55	0.53	0.49	0.55	78
		MAE	0.38	0.42	0.45	0.43	0.41	0.45	
		R	0.46	0.34	0.04	0.12	0.31	0.06	
	Daegu	RMSE	0.42	0.62	0.64	0.70	0.63	0.67	83
		MAE	0.34	0.50	0.54	0.60	0.54	0.56	
		R	0.79	0.34	0.10	-0.19	0.17	0.10	
	Gwangju	RMSE	0.54	0.76	0.94	0.84	0.79	0.66	88
		MAE	0.39	0.54	0.74	0.70	0.67	0.56	
		R	0.61	0.24	-0.28	-0.16	0.06	0.61	
Wet (SPI3 ≥ 1)	Seoul	RMSE	0.73	1.23	1.45	1.55	1.53	1.62	16
		MAE	0.65	1.11	1.36	1.45	1.43	1.51	
		R	0.43	0.15	0.11	-0.21	-0.16	-0.25	
	Daejeon	RMSE	1.25	1.39	1.40	1.43	1.40	1.55	20
		MAE	1.22	1.36	1.37	1.40	1.37	1.52	
		R	0.34	0.29	0.34	0.13	0.04	0.09	
	Daegu	RMSE	0.68	1.04	1.44	1.49	1.31	1.33	19
		MAE	0.58	0.99	1.37	1.42	1.26	1.22	
		R	0.63	0.67	0.28	0.27	0.58	-0.02	
	Gwangju	RMSE	0.45	0.84	1.34	1.20	0.94	0.70	21
		MAE	0.40	0.77	1.24	1.11	0.87	0.63	
		R	0.43	0.16	-0.30	-0.29	0.05	0.24	

Table 6. Comparison of Forecasting Results between Flood and Drought Season

Forecasting Measure		Season	Flood Season (Jun, Jul, Aug, Sep)						Lead Time (months)	Drought Season (Oct, Nov, Dec, Jan, Feb, Mar, Apr, May)					
			1	2	3	4	5	6		1	2	3	4	5	6
		Season							Season						
RMSE	SPI3	Entire	0.75	0.86	0.91	0.91	0.86	0.86	Entire	0.75	0.86	0.91	0.91	0.86	0.86
		Flood	0.71	0.93	0.97	0.90	0.94	0.88	Drought	0.77	0.81	0.87	0.92	0.82	0.85
	SPI6	Entire	0.62	0.77	0.82	0.81	0.79	0.78	Entire	0.62	0.77	0.82	0.81	0.79	0.78
		Flood	0.48	0.73	0.78	0.75	0.68	0.66	Drought	0.34	0.53	0.61	0.60	0.59	0.59
	SPI12	Entire	0.52	0.63	0.67	0.69	0.71	0.74	Entire	0.52	0.63	0.67	0.69	0.71	0.74
		Flood	0.80	0.88	0.88	0.85	0.81	0.79	Drought	0.29	0.46	0.52	0.60	0.65	0.71
MAE	SPI3	Entire	0.58	0.66	0.68	0.67	0.65	0.72	Entire	0.58	0.66	0.68	0.67	0.65	0.72
		Flood	0.53	0.74	0.74	0.70	0.73	0.72	Drought	0.60	0.63	0.65	0.66	0.61	0.73
	SPI6	Entire	0.47	0.60	0.64	0.62	0.61	0.63	Entire	0.47	0.60	0.64	0.62	0.61	0.63
		Flood	0.52	0.67	0.69	0.66	0.64	0.63	Drought	0.44	0.57	0.61	0.59	0.59	0.63
	SPI12	Entire	0.35	0.43	0.49	0.52	0.55	0.59	Entire	0.35	0.43	0.49	0.52	0.55	0.59
		Flood	0.61	0.68	0.70	0.66	0.63	0.65	Drought	0.06	0.06	0.07	0.07	0.09	0.10
R	SPI3	Entire	0.58	0.26	0.04	-0.08	0.10	0.15	Entire	0.58	0.26	0.04	-0.08	0.10	0.15
		Flood	0.60	0.15	0.04	0.08	0.01	0.19	Drought	0.54	0.31	0.16	-0.14	0.19	0.05
	SPI6	Entire	0.62	0.32	0.21	0.15	0.16	0.22	Entire	0.62	0.32	0.21	0.15	0.16	0.22
		Flood	0.48	0.42	0.50	0.42	0.52	0.30	Drought	0.68	0.41	0.32	0.26	0.16	0.24
	SPI12	Entire	0.77	0.65	0.61	0.57	0.56	0.52	Entire	0.77	0.65	0.61	0.57	0.56	0.52
		Flood	0.55	0.50	0.49	0.47	0.50	0.43	Drought	0.93	0.83	0.78	0.71	0.65	0.55

한 결과 RMSE와 MAE의 경우에는 SPI 3, 6, 12 모두 홍수기에 대한 SPI를 예측한 결과에 비해서 갈수기에 대한 SPI 예측 결과의 오차가 매우 적은 것으로 나타났다. 반대로 R의 경우에는 모든 선행예보시간에 대해서 갈수기가 홍수기보다 크게 나타났다. 이를 통하여 다층 퍼셉트론 인공지능망 모형과 SPI를 이용한 가뭄예측 홍수기간에 비하여 가뭄기간의 예측결과가 매우 높은 것으로 나타났다. 인공지능망 모형이 가뭄예측에 매우 효율적인 것으로 분석되었다.

### 5. 결 론

본 연구에서는 인공지능망 다층 퍼셉트론을 이용하여, 남한 유역내 주요 지점을 대상으로 과거 관측 자료를 통해 산정한 서로 다른 지속기간을 갖는 비정상성 SPI 시계열자료를 이용하여 가뭄예측 시도하였으며, 예측된 결과

를 관측치와 비교하여 다음과 같은 결론을 도출하였다.

- 1) 인공지능망 다층퍼셉트론을 적용하여 DMSNN방법을 이용한 학습은 입력층과 출력층에 변수를 6개월에서 12개월까지 증가하고, 은닉층에 개수를 제한하는 방법으로 학습을 실시한 결과 입력층과 출력층에 변수를 12개월로 하여, 은닉층을 12~16개로 제한하였을 때 산정한 구조가 가뭄예측을 위해 적합한 구조로 나타났다.
- 2) RMSE의 경우 주요 4개 관측소 모두 선행 예보 기간이 1~6개월로 길어질수록 오차가 증가하는 것으로 나타났으며, 지속기간이 길어질수록 오차가 감소하는 것으로 분석되었다.
- 3) MAE의 경우 RMSE와 같은 추세를 보여주고 있다. 이는 예측기간이 길어질수록 가뭄예측에 대한 오차가 증가하는 것으로 판단되며, 지속기간이 길수록 가뭄예측에 신뢰도가 높은 것으로 나타났다.

- 4) R의 경우 RMSE, MAE와 반대로 주요 4개 관측소 모두 예측기간이 1~6개월로 기간이 길어질수록 상관성이 낮은 것으로 나타났으며, 지속기간이 길수록 상관성이 높은 것으로 분석되었다.
- 5) 홍수기와 갈수기를 분리하여 결과를 비교한 결과, 다층 퍼셉트론 인공신경망 모형을 적용한 결과 RMSE, MAE에서 갈수기에 오차가 적으며, R에서 갈수기에 상관성이 홍수기보다 높게 나타났으며, 이는 가뭄을 예측하기에 다층퍼셉트론 인공신경망이 적합한 것으로 판단된다.
- 6) 다층퍼셉트론을 이용한 가뭄예측 결과 SPI (3)의 경우 선행예보 1개월이 가뭄예측이 잘 되었고, 1개월 이후로 정확도가 감소하는 경향이 나타나, 단기 가뭄 예측에 활용하는 것이 좋을 것으로 사료된다. SPI (6)의 경우 1~2개월까지 가뭄예측이 잘 이루어 졌고, 2개월 이후에서 예측결과에 대한 신뢰도가 급격하게 감소하는 것으로 나타났다. 이를 통해 중기 가뭄예측에 활용하는 것이 좋을 것으로 판단된다. SPI (12)의 경우 1~5개월까지 가뭄예측이 다른 SPI (3), SPI (6)보다 높은 예측치를 보여주는 것으로 나타났다.

이상의 연구를 바탕으로 남한 유역내 주요 지점을 대상으로 가뭄을 예측한 결과 주요 4개 관측소 모두 과거 가뭄 (2001~2010)기간과 학습을 통해 산정한 구조로 가뭄 예측한 결과 가뭄이 잘 예측된 것으로 나타났으며, 갈수기와 홍수기를 분류한 결과를 통해 다층 퍼셉트론이 가뭄예측에 적합한 것으로 분석되었으며, 다층퍼셉트론을 이용한 가뭄예측 결과를 통해 SPI (3), (6), (12) 모두 단기간의 Normal 등급에 가뭄 예측에 적합한 것으로 판단되었다.

## 감사의 글

본 연구는 기상청 기상기술개발사업(CATER 2012-3100, 한반도/동아시아 가뭄전망·대응기술 개발)의 지원으로 수행되었으며 이에 감사드립니다.

## References

- Bacanal, U.G., Firat, M., and Dikbas, F. (2009). "Adaptive Neuro-Fuzzy Inference System for drought forecasting." *Stochastic Environmental Research and Risk Assessment*, Vol. 23, No. 8, pp. 1143-1154.
- Belayneh, A., and Adamowski, J. (2012). "Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression." *Applied computational Intelligence and Soft Computing*, Vol. 2012 No. 6, DOI=10.1155/2012/794061.
- Chen, J., Huang, Z., and Jin, Q. (2012). "SPI-based drought characteristics analysis and prediction for Xiqiao Station in Yunnan Province, China." *Disaster Advances*, Vol 5, pp. 396-407
- Ghosh, S., and Mujumdar, P.P. (2007). "Nonparametric methods for modeling GCM and scenario uncertainty in drought assessment." *Water Resources Research*, Vol. 43, No 7, W07405-W07406.
- Jeong, H.J., Lee, S.J., and Lee, H.K. (2002). "Water Quality Forecasting of Chungju Lake Using Artificial Neural network Algorithm." *Korean Environmental Science Society*, Vol. 11, No. 3, pp. 201-207.
- Kang, B.S., and Lee, B.K. (2011). "Application of Artificial Neural Network to Improve Quantitative Precipitation Forecasts of Meso-scale Numerical Weather Prediction." *Journal of Korea Water Resources Association*, Vol. 44, No. 2, pp. 97-107.
- Kim, C.J., Park, M.J., and Lee, J.H. (2013). "Analysis of climate change impacts on the spatial and frequency patterns of drought using a potential drought hazard mapping approach." *International Journal of Climatology*, *Published in Online*, DOI=10.1002/joc.3666.
- Kwon, H.H., Moon, J.W., Song, H.S., and Moon, Y.I. (2009). "Climate Information and GCMs Seasonal Forecasts based Short-term Forecasts for Drought." *Conference of Korea Water Resources Association*, pp. 1186-1190.
- Mckee, T.B., Doesken, N.J., and Kleist, J. (1993). *The relationship of drought frequency and duration of time scales*. 8th Conference on Applied Climatology, Jan., Anaheim, CA, pp. 179-184.
- Mckee, T.B., Doesken, N.J., and Kleist, J. (1995). Drought monitoring with multiple time scales preprints. 9th Conference on Applied Climatology, 15-20 January, Dallas, TX, pp. 233-236.
- Mishra, A.K., and Desai, V.R. (2006). "Drought forecasting using feed-forward recursive neural network." *Ecological Modelling*, Vol. 198, Issue 1-2, pp. 127-138.
- Morid, S., Smakhtin, V., and Bagherzadeh, K. (2007).

“Drought forecasting using artificial neural networks and time series of drought indices.” *International Journal of Climatology*, Vol. 27, No. 15, pp. 2103–2111.

Paulo, A.A., Ferreira, E., Coelho, C., and Pereira, L.S. (2005). “Drought class transition analysis through Markov and Loglinear models, an approach to early warning.” *Agricultural Water Management*, Vol. 77, pp. 59–81.

Seo, J.W. (2011). *Analysis on the Statistical Characteristics of Drought in Korea using SPI and PDSI*.

M.S. dissertation. University of Kyung Hee, Seoul, pp. 44

Souhaib B.T., and Rob, J.H. (2012). *Recursive and Direct multi-step forecasting: the best of both worlds*. Ph.D. dissertation, University of Monash, Victoria, Clayton, Australia, pp. 19–12.

논문번호: 13-070	접수: 2013.09.01
수정일자: 2013.10.21/11.06/12.02	심사완료: 2013.12.02