

# 인공지능기법을 이용한 일유출량의 추계학적 비선형해석

## A Stochastic Nonlinear Analysis of Daily Runoff Discharge Using Artificial Intelligence Technique

안 승 섭\* · 김 성 원\*\*  
Ahn, Seung Seop · Kim, Sung Woon

### Summary

The objectives of this study is to introduce and apply neural network theory to real hydrologic systems for stochastic nonlinear predicting of daily runoff discharge in the river catchment. Back propagation algorithm of neural network model is applied for the estimation of daily stochastic runoff discharge using historical daily rainfall and observed runoff discharge. For the fitness and efficiency analysis of models, the statistical analysis is carried out between observed discharge and predicted discharge in the chosen runoff periods. As the result of statistical analysis, method 3 which has much processing elements of input layer is more prominent model than other models(method 1, method 2) in this study.

Therefore, on the basis of this study, further research activities are needed for the development of neural network algorithm for the flood prediction including real-time forecasting and for the optimal operation system of dams and so forth.

### I. 서 론

하천유역에서 수자원 관리 및 운영을 위한 계획에서 발생하는 문제의 해결을 위한 기본적인이고 중요한 요소중의 하나는 하천유출량의 정확한 추정이라 할 수 있다. 하천의 유출량은 주로 강우가 하천유역이라는 복잡한 시스

템을 통하여 발생하는 결과로 볼 수 있다. 하천유역은 수문기상학적인자를 포함하는 시스템으로 간주할 수 있으며, 여기에서 발생하는 강우-유출현상은 비정규성과 불확실성을 내포하기 때문에 강한 비선형성을 나타낸다. 따라서 정확한 하천유출량의 추계학적 예측은 얼마만큼 정확히 시스템을 이용하는가에 달려있

\*경일대학교 공과대학  
\*\*경일대학교 공과대학(강사)

키워드 : 신경망이론, 역전파알고리즘, 일유출량, 누가밀도함수, 통계분석, 1차 계열상관계수, 계열상관도, 비선형해석

다. 수학적모형으로 개발된 신경망(neural network)은 함수가 가지고 있는 비선형학습 및 보간능력으로 인하여 응용분야에서 광범위하게 이용되고 있는 실정이다.

그러므로 신경망이론은 과거자료의 입출력 패턴으로부터 정보를 추출하여 지식으로 보유하고 이를 근거로 새로운 상황에 대한 해답을 제시하도록 하는 인공지능분야의 학습기법이다. 인간이 경험과 훈련에 의하여 지식을 축적해 가듯이 시스템의 입력과 출력에 의하여 연결강도를 최적화함으로써 모델의 구조를 스스로 조직화하기 때문에 유출계의 모형과정(model process)을 설정하지 않더라도 입출력에 따라 모델의 구조에 적합한 매개변수를 최적 추정할 수 있다. 따라서 수학적 알고리즘의 적용이 어려운 강우와 유출관계를 하천유역이라는 시스템에서 발생한 신호체계의 입출력패턴으로 간주하고 인간의 사고과정을 근거로 신경망의 수학적 알고리즘을 파악하여 수문계의 예측에 적용할 수 있을 것이다.

McCulloch와 Pitts(1943)에 의하여 제안된 신경망이론을 이용한 강우-유출해석에 관한 연구로는 외국의 경우 Karunanithi & Grenney(1994)<sup>8)</sup>는 하천유출예측에 있어서 신경망을 적용하여 분석하였으며, Lepedes & Farber(1987)<sup>9)</sup>는 신경망이론을 시계열 예측에 처음 이용하였는데, 이는 시계열자료를 다층인식자에 적용하여 오차역전파(EBP : Error Back Propagation)알고리즘에 의해 학습시킨 것으로 관측치에 대한 적합도(Fitting) 및 예측치가 기존의 계량적 방법보다 우수하다고 하였다. 그리고 Jason & Robert(1995)<sup>7)</sup>는 지류를 포함한 복합하천유역에 신경망모델을 적용하여 유출과정을 시뮬레이션시켜서 그 타당성을 검토하였으며, Smith(1992)<sup>10)</sup>는 백프로파게이션 신경망모델을 이용하여 하천유역의 홍수유출특성을 분석하여 예측모델을 구성하였다. 또한 국내의 연구로는 김주환

(1993)<sup>2)</sup>은 신경망의 복잡한 시스템 등정과정 없이 학습에 의하여 연결강도를 조절함으로써 유출예측모형을 구성하였으며, 오남선(1996)<sup>3)</sup>은 신경망이론을 강우예측모델에 적용하여 서울 및 소양강유역의 강우자료를 이용하여서 그 결과를 분석한 바 있으며, 이남호(1994)<sup>4)</sup>등은 백프로파게이션과 카운터프로파게이션 알고리즘을 적용하여 신경회로망을 이용한 수도증발산량 예측에 이용하여 예측이론의 적용가능성을 제시한바 있다. 그러나 아직까지는 신경망이론이 강우-유출해석에 도입되어 진행된 연구성과는 외국에 비해 극히 부진한 실정에 있다.

따라서 본 연구에서는 강우-유출을 유역이라는 시스템내에서 발생한 신호체계의 패턴으로 간주하고 모델을 구성하는 각 유니트간의 매개변수를 인공신경망 모델의 역전파알고리즘으로 최적화하는 방법을 실제수문계에 적용하고 홍수기의 일단위 하천유출량의 과거기록치를 이용하여 유출량을 추계학적으로 예측하고자 한다. 또 그 결과를 비교분석하여 수문량예측의 적용가능성을 제시하고자 한다.

## II. 신경망모델의 기본이론

그물과 같이 연결된 뇌의 정보처리과정의 동적거동을 표현하기 위하여 개발된 인공신경망은 정보, 통신, 제어분야 등 복잡한 문제에 적용되고 있다. 인간의 두뇌를 형성하고 있는 신경조직은 뉴런(neuron)이라는 신경소자가 복잡하게 엮힌 비선형시스템으로 이들은 서로 연결해 주는 결합강도에 의하여 시스템의 성능을 향상시킨다. 각 뉴런은 다른 뉴런으로부터 신호를 입력받아 이를 종합하여 뉴런의 반응여부인 출력을 결정한다. 반응을 할 경우 뉴런은 전기펄스를 내어서 연결된 다른 뉴런에 신호로서 전달한다. 이러한 입력과 출력의 상호작용은 동적과정으로 표현 될 수 있으며

입력과 출력간의 상관관계는 뉴런간의 결합강도에 의존한다. 또한 결합강도는 학습과 자기조직에 의하여 조정된다. 신경회로망은 이러한 뇌와 동일한 계산방식을 사용한 커넥셔니즘(connectionism)에 기반을 둔 수리모형이다.

신경망모델은 학습의 방법에 따라 미리 훈련할 값을 신경회로망에 제시해주고 그 목표에 도달할 때까지 회로망의 연결강도를 조정해나가는 지도학습(supervised learning)모델과 스스로의 조직화를 통하여 특징들을 추출하여 학습하는 자율학습(unsupervised learning)모델로 크게 나눌수 있다. 따라서 지도학습에서 신경망은 출력값과 비교할 수 있는 기지의 정답이 존재하며, 그 오차를 이용하여 학습을 수행하게 된다. 이와 반대로 자율학습에서는 이용가능한 정보로 자료나 신호(signal)만이 제공되며 신경망은 스스로 입력자료의 흥미있는 범주나 특징을 발견하여야 하므로 정확한 정보없이 자기조직화에 의해 훈련을 하게 되고 결과를 출력하게 된다. 또한 신경망은 일반적으로 입력자료와 출력자료가 이진수(binary)의 형태를 가지거나 실수형태의 연속적인 값으로 나타나는 두가지로 나뉘어진다. 신경망의 학습에는 자료를 정규화(normalization)하는 전처리과정이 필요하며, 이것은 처리소자들의 집합으로 되어있는 신경망의 효율을 높이기 위하여 잡음이나 왜곡에 의해 변형된 것을 바로잡는 기초적인 연산처리과정이며, 신경망에서 출력된 값을 다시 역변환하여 처리될 후처리에 대응한다. 그리고 처리소자의 시그모이드 전달함수(Sigmoid transfer function)는 최소값 0, 최대값 1을 출력한다. 따라서 수문과정의 입출력자료를 신경망에 적용하기 위한 전처리 과정에는 적용자료의 최대값과 최소값을 결정하여 0과 1사이에 분포시키거나 이진수의 형태로 표현하여야 한다.

본 연구에서는 기왕의 자료를 이용하는 지도학습을 이용한 역전파 알고리즘을 사용하

고, 학습을 위한 입력자료로 이용되는 강우-유출량자료가 상수형태의 연속적인 값으로 존재하며 이를 바탕으로 신경망을 구성하고자 한다. 또한 유출량예측을 위하여 실수형태의 자료에 대하여 지도학습을 실시할 수 있으며 비선형함수를 모의할 수 있는 다층신경망모델을 구성하였다. 다층신경망모델에서 은닉층의 개수가 1개일 경우 Fig. 1과 같은 이층신경망모델을 구성할 수 있으며, Fig. 2는 Fig. 1에서의 은닉층과 출력층의 각 뉴런의 입·출력관계를 도시한 것이다.

Fig. 1에서  $O_i$ 는 출력층(output layer)성분,  $V_j$ 는 은닉층(hidden layer)성분,  $\xi_k$ 는 입력층(input layer)성분,  $W_{jk}$ 와  $W_{ij}$ 는 입력층에서 은닉층으로의 연결강도와 은닉층에서 출력층으로의 연결강도를 각각 표시하였다. 입력자료의 서로 다른 패턴을  $\mu$ 로 구분하면, 입력자료 가운데 k번째 성분은  $\xi_k^\mu$ 로 표기되며 이 경우 은닉층의 j번째 성분으로의 입력  $h_j^\mu$ 은 식 (1)과 같이 계산할 수 있으며, 은닉층의 j번째 성분  $V_j^\mu$ 은 시그모이드 전달함수  $g(x)$ 를 통하여 식 (2)와 같이 나타낼 수 있다.

이와같은 원리에 의해 출력층의 i번째 성분

$$h_j^\mu = \sum_k W_{jk} \xi_k^\mu \dots\dots\dots (1)$$

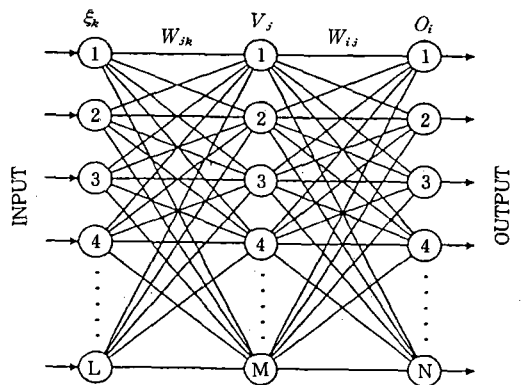


Fig. 1. Structure of two-layer neural network model

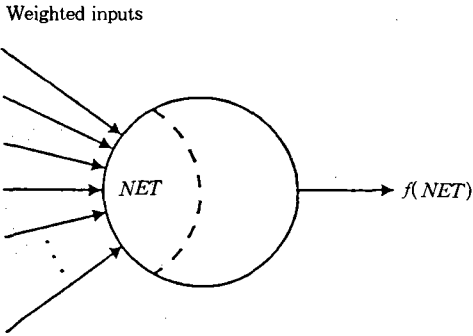


Fig. 2. Input-output relationship of each neuron in layers j and i of Fig. 1

$$V_j^\mu = g_1(h_j^\mu) = g_1(\sum_k W_{jk} \xi_k^\mu) \dots\dots\dots (2)$$

으로의 입력성분은 식 (3)과 같이 정의할 수 있으며, 이에 따른 출력성분은 시그모이드 전달함수  $g_2(x)$ 를 이용하여 식 (4)와 같이 나타낼 수 있다.

$$h_i^\mu = \sum_j W_{ij} V_j^\mu = \sum_j W_{ij} g_1(\sum_k W_{jk} \xi_k^\mu) \dots\dots (3)$$

$$O_i^\mu = g_2(h_i^\mu) = g_2(\sum_j W_{ij} V_j^\mu) \\ = g_2(\sum_j W_{ij} g_1(\sum_k W_{jk} \xi_k^\mu)) \dots\dots\dots (4)$$

따라서 식 (4)는 출력패턴  $O^\mu$ 과 입력패턴  $\xi^\mu$ 의 관계를 나타내는 신경망의 구조로서 이에 필요한 연결강도  $W_{jk}$ 와  $W_{ij}$ 를 기왕의 자료로부터 학습을 통하여 적절하게 조절해 나가게 된다.

Hertz(1991)<sup>6)</sup>등에 의하면, 이러한 다층구조의 신경망은 뛰어난 적용성에도 불구하고 학습방법이 없어 활용되지 못하다가 1980년대 중반 역전파(back propagation) 알고리즘을 이용한 다층구조의 학습방법이 발표된 이후 이를 이용한 신경망에 관한 연구가 활성화되기 시작했다. 이 알고리즘은 입력층의 각 처리요소에 입력패턴을 주면 신호가 변환되어 은닉층에 전달되고 출력층에서 이 신호를 출

력시킨 후 출력치와 기대치를 비교하여 오차가 발생하면, 오차를 감소시키도록 연결강도를 조정하여 하위층으로 역전파하여 다시 연결강도를 조절하는 원리를 이용한다. 즉, 주어진 입력과 출력에 대한 적절한 연결강도를 구하기 위하여 식 (5)와 같이 신경회로망의 출력치( $O_i^\mu$ )와 기대출력치( $\zeta_i^\mu$ )사이의 오차의 제곱을 최소화하는 비용함수  $E(W)$ 의 형태로 정의한다.

$$E(W) = -\sum_{\mu} (\zeta_i^\mu - O_i^\mu)^2 \dots\dots\dots (5)$$

그리고, 이 알고리즘에서는 학습이 진행됨에 따라 정의된 비용함수가 0에 가까워지게 되므로서 적절한 신경망이 완성되며, 이를 위하여 경사하강(Gradient Decent)알고리즘을 이용하게 되고 이에 따른 연결강도들의 조절량은 다음 식 (6) 및 (7)과 같다.

$$\Delta W_{ij} = \eta \sum_{\mu} \delta_i^\mu V_j^\mu \dots\dots\dots (6)$$

$$\Delta W_{jk} = \eta \sum_{\mu} \delta_j^\mu \xi_k^\mu \dots\dots\dots (7)$$

여기서,  $\eta$ 는 학습율(learning rate)이고,  $\delta_i^\mu$ 와  $\delta_j^\mu$ 는 식 (8)~(9)와 같이 전달함수의 미분 형태와 오차의 곱으로 나타낼 수 있다.

$$\delta_i^\mu = g'(h_i^\mu) [\zeta_i^\mu - O_i^\mu] \dots\dots\dots (8)$$

$$\delta_j^\mu = g'(h_j^\mu) W_{jk} \delta_i^\mu \dots\dots\dots (9)$$

식 (9)에서 입력층과 은닉층사이에 연결강도조절에 필요한  $\delta$ 를 계산하기 위하여 출력층  $O_i$ 로부터 계산된  $\delta$ 값들이 필요하다. 즉, 입력 신호가 신경망내에서 입력층으로부터 출력층으로 전파해 가는데 비해 오차의 값들은 역방향으로 전파해 나가기 때문에 식 (6)과 (7)에 근거하여 신경망의 연결강도를 조절해 나가

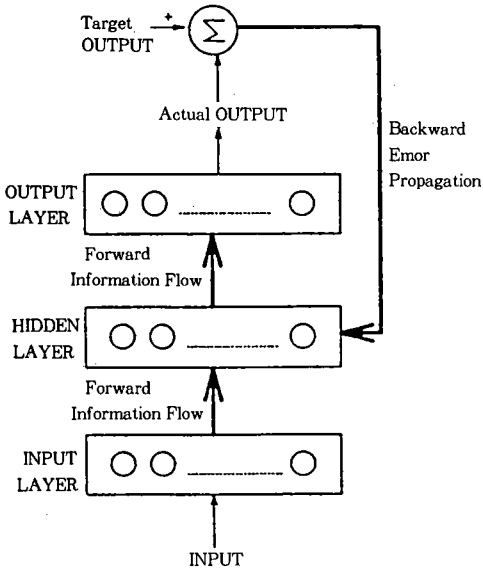


Fig. 3. Schematic diagram of back-propagation network

방법을 역전파 알고리즘이라고 하며, Fig. 3과 같이 도식화할 수 있다.(이남호 등, 1994; 오남선 등, 1996)

### III. 일유출량 예측모델의 구성

#### 1. 유출모델의 구성

본 연구에서는 하천유역의 일유출량 예측모델을 구성하기 위한 입력층의 성분으로 일강우량자료 및 일유출량자료를 이용하였고, 출력층의 성분으로 예측 일유출량이 유도되며 이것은 실수형태의 연속적인 값이다. 따라서 신경망을 이용한 일유출량 예측을 위하여 다층신경망모델을 이용하였고, 다층신경망의 구조는 기존의 연구에서 주로 이용한 것과 같이 입력층과 출력층사이에 하나의 은닉층이 존재하는 이층신경망을 이용한다. 본 연구에서는 유출량의 예측을 위한 신경망의 구조를 시그모이드 전달함수에 의한 시계열 모형으로 가정하였다. 이 관계를 시그모이드 전달함수의

입출력변수를 이용하면 다음 식 (10)과 같이 나타낼 수 있다.

$$Q_k = \Gamma(Q_{k-1}, Q_{k-2}, \dots, Q_{k-nh}, P_{k-1}, P_{k-2}, \dots, P_{k-nh}) \dots \dots \dots (10)$$

여기서 전달함수모델의 입출력변수가 각각 신경회로망의 입력층, 출력층 처리소자가 되며 신경망을 비선형함수  $\Gamma(\ )$ 로 표현하여 그 관계를 나타낼 수 있다. 따라서 신경망을 미지의 비선형함수  $\Gamma(\ )$ 로 표현하였고 함수의 괄호내 입력변수는 전달함수모델의 입력변수로서 과거의 일강우량 및 일유출량자료이다. 그리고 이에 대응하는 출력패턴으로 예측유출량  $Q_k$ 가 출력되도록 각 모델을 다음 식(10.a)~(10.c)와 같이 가정하였다.

Method 1

$$Q_k = \Gamma(Q_{k-1}, Q_{k-2}, Q_{k-3}, P_{k-1}, P_{k-2}) \dots \dots \dots (10.a)$$

Method 2

$$Q_k = \Gamma(Q_{k-1}, Q_{k-2}, Q_{k-3}, Q_{k-4}, P_{k-1}, P_{k-2}, P_{k-3}) \dots \dots \dots (10.b)$$

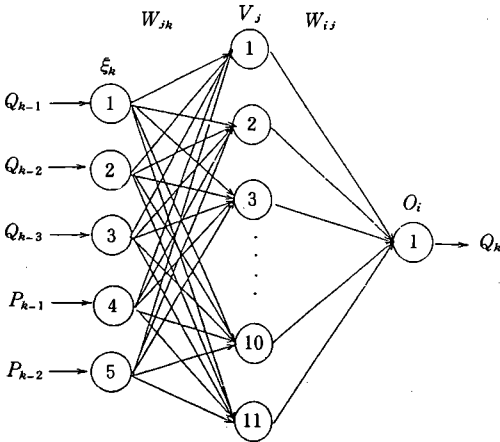
Method 3

$$Q_k = \Gamma(Q_{k-1}, Q_{k-2}, Q_{k-3}, Q_{k-4}, Q_{k-5}, P_{k-1}, P_{k-2}, P_{k-3}, P_{k-4}, P_{k-5}) \dots \dots \dots (10.c)$$

입출력 변수에 대한 미지의 비선형함수는 신경망으로부터 추정된 연결강도로서 예측모델식을 구성하는 매개변수의 역할을 하며, 신경회로망의 구성은 식 (10.a)~(10.c)에서 입력변수에 대한 입력층 처리소자와 bias 그리고 입력층 처리소자수의 2d 또는 2d+1개의 중간층의 처리소자, 출력층 처리소자는 1단계 전 예측값이 출력되도록 구성하였다. 또한 일유출량 해석을 위한 계산과정은 학습을 통하여 추정된 매개변수로서 각 층간의 연결강도를 이용하여 출력층에서 예측유출량을 출력하

**Table 1. Number of processing element for flood prediction models**

Index	Input layer (I)	Hidden layer (J)	Output layer (K)
Method 1	5	11	1
Method 2	7	15	1
Method 3	10	21	1



**Fig. 4. Neural network architecture of daily-runoff prediction model(Method 1)**

게 된다.

각 모델에 따라서 1990년이후 1991년, 1992년, 1994년의 홍수기인 6월에서 9월사이의 실측 일유출량을 학습자료(training data)의 입력데이터로 사용하였으며 1995년의 홍수기인 6월부터 9월사이의 실측 일유출량을 검증자료(checking data)로 사용하였다. 또한 각 층의 연결강도의 초기치는 어떠한 값을 사용해도 매개변수가 최적해로 수렴하게 되므로 본 연구에서는 -1~1사이의 난수를 발생시켜 초기화하였고, 학습률  $\eta$ 와 모멘텀상수  $\alpha$ 는 Yoh-Han Pao(1989)<sup>11)</sup>에 의하여 실험적으로 제한한 학습률 0.9, 모멘텀상수 0.7을 이용하여 유출예측에 사용하였으며, 분산처리를 통한 효율적인 학습을 위하여 입력층 및 은닉층에 두는 bias는 1로 가정하였다. Table 1은 일유출량예측을 위한 신경망의 입력층, 은닉층

및 출력층 처리소자의 수를 나타낸 것이고, Fig. 4는 일유출량 예측모델중 Method 1에서의 신경망구조를 나타낸 것이다.

## 2. 유출모델의 통계분석

본 연구에서는 유출량 예측문제에 적용되고 있는 통계적 검증방법으로 각 모델에서 계산된 오차와 비교검토하여 신경망모델의 적합도 및 정확도를 판단하였다.

관측유량  $Q_i$ 와 모델로부터 예측된 유량  $\hat{Q}_i$  사이에 예측오차  $E_i$ 는 다음 식 (11)과 같이 나타낼 수 있다.

$$E_i = Q_i - \hat{Q}_i \dots\dots\dots (11)$$

오차의 크기는 일정기간 동안에 발생하는 모든 예측오차의 합계를 고려해야 한다. 그러나  $\sum(Q_i - \hat{Q}_i)$ 은 오차의 크기에 관계없이 합계가 거의 0이 되므로 이 경우에는 예측오차의 절대값을 검토하게 된다. 즉 절대편차(absolute deviation)는  $|E_i| = |Q_i - \hat{Q}_i|$ 이므로 이 절대편차를 자료의 크기 N으로 나누어 평균절대편차(mean absolute deviation : MAD)를 구하며, 다음 식 (12)와 같이 나타낼 수 있다.

$$MAD = \frac{1}{N} \sum_{i=1}^N |E_i| = \frac{1}{N} \sum_{i=1}^N |Q_i - \hat{Q}_i| \dots (12)$$

또 평균절대편차에서 음(-)과 양(+의) 예측오차가 상쇄되는 것을 예방할 수 있는 또 다른 방법은 예측오차를 제곱하는 제곱오차(square errors)  $(E_i)^2 = (Q_i - \hat{Q}_i)^2$ 으로부터 평균제곱오차(mean square errors : MSE)를 구하면 다음 식 (13)과 같이 나타낼 수 있다.

$$MSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i)^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_i - \hat{Q}_i)^2} \dots\dots\dots (13)$$

상대적인 오차의 크기를 나타내는 평균절대 백분비오차(mean absolute percentage errors ; MAPE)와 시계열오차(time series error ; TSE)는 다음 식 (14) 및 (15)와 같다.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|Q_i - \hat{Q}_i|}{Q_i} \times 100(\%) \dots (14)$$

$$TSE = \frac{1}{N} \sum_{i=1}^N \frac{Q_i - \hat{Q}_i}{Q_i} \times 100(\%) \dots (15)$$

예측결과에 대한 정확성 및 적합성은 위 식으로 계산한 절대오차, 상대오차값을 구하여 이들의 값이 적을수록 적합성이 좋은 것이 된다. 또한 모델의 효용도(model efficiency)는 다음 식 (16)과 같이 결정계수( $r^2$ : determination coefficient)를 계산하여  $r^2$ 의 값이 1에 가까울수록 정확도와 적합성이 양호한 것으로 한다.

$$r^2 = 1 - \frac{\sum_{i=1}^N (Q_i - \hat{Q}_i)^2}{\sum_{i=1}^N (Q_i - \bar{Q}_i)^2} \dots (16)$$

여기서,  $\bar{Q}_i$ 는 유량자료계열의 평균치이다.

한편 상대오차의 평균을 나타내는 Bias 및 체적오차의 백분율을 나타내는 VER(volume error)는 식 (17) 및 (18)과 같이 나타낼 수 있으며, 그 값이 0에 가까울수록 정확도와 적합성이 양호한 것으로 한다.

$$Bias = \frac{1}{N} \sum_{i=1}^N (Q_i - \hat{Q}_i) \dots (17)$$

$$VER = \frac{\sum_{i=1}^N Q_i - \sum_{i=1}^N \hat{Q}_i}{\sum_{i=1}^N Q_i} \times 100 \dots (18)$$

### 3. 유량계열의 추계학적해석

가. 1차 계열상관계수에 의한 수문학적 지속성의 판정

본 연구에서는 유량자료계열의 추계학적 분석을 위하여 수문학적 지속성을 판정하는 기준으로 1차 계열상관계수(lag-one serial correlation coefficient)에 의한 분석을 실시하였으며, 유량계열  $Q_i$ 의 1차 계열상관계수는 다음 식 (19)와 같이 나타낼 수 있다.

$$r_1 = \frac{C_1}{C_0} \dots (19)$$

여기서,  $C_0$ 와  $C_1$ 은 각각 영차 혹은 1차 계열공분산(serial covariance)으로서 자료수  $n$ 이 클 때에는 다음 식 (20) 및 (21)과 같이 나타낼 수 있다.

$$C_0 = \frac{1}{n} \sum_{i=1}^n (Q_i - \bar{Q}_i)^2 \dots (20)$$

$$C_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} (Q_i - \bar{Q}_i)(Q_{i+1} - \bar{Q}_i) \dots (21)$$

식 (19)로 나타내는 1차 계열상관계수는 자료계열을 구성하는 각 유량치가 한 시간단위만큼 선행하는 유량치와 얼마만한 수문학적 지속성을 가지는 것을 나타내며 완전 무작위 시계열의 경우는 이론적으로는  $r_1=0$ 이고, 지속성이 강한 시계열의 경우에는  $r_1$ 이 1에 수렴하게 된다. 1차 계열상관계수에 의한 자료계열의 수문학적 지속성 판정법에는 몇가지가 존재하나, 본 연구에서는 Anderson, R. L. (1941)<sup>5)</sup>에 의해 제시된 exact test를 이용하여 분석하고자 한다. 따라서  $r_1$ 은 평균치  $\mu$ 와 분산  $\sigma^2$ 의 함수로 나타낼 수 있으며, 다음 식 (22)와 같이 나타낼 수 있다.

$$r_1 = \mu + \sigma z = -\frac{1}{n-1} + z \frac{(n-2)}{(n-1)^{3/2}} \dots (22)$$

식 (22)에서 95% 신뢰구간은  $F(z)=0.025$ 와 0.975일 때의  $z = \pm 1.96$ 사이이므로  $r_1$ 의 범위는 다음 식 (23)과 같이 나타낼 수 있다.

$$-\frac{1}{n-1} - 1.96 \frac{(n-2)}{(n-1)^{3/2}}$$

$$< r_1 < -\frac{1}{n-1} + 1.96 \frac{(n-2)}{(n-1)^{3/2}} \quad \dots (23)$$

유량계열로부터 계산된  $r_1$ 이 식 (23)과 같이 신뢰구간(confidence level)내에 들어가면 수문학적 지속성이 거의 없는 무작위 계열로 간주하고, 신뢰구간 바깥에 위치하면 연속성이 있는 계열로 분석된다.

나. 계열상관도에 의한 수문학적 지속성의 판정

본 연구에서는 유량자료계열의 수문학적 지속성을 판별하기 위하여 자료의 계열상관도를 작성하여 분석함으로써 판단할 수 있다. 계열상관도는 계열상관계수의 차수(number of lags)와 상관계수  $r_k$ 간의 관계를 표시하는 도면으로서,  $k$ 차 계열상관계수는 어떤 시점에서의 유량이  $k$ 시간단계후의 유량과 가지는 상관도를 표시하는 것으로 다음 식 (24)와 같이 나타낼 수 있다.

$$r_k = \frac{C_k}{C_0} \dots\dots\dots (24)$$

여기서,  $C_0$ 는 전절의 식 (20)과 같고,  $C_k$ 는 식 (25)와 같이 나타낼 수 있다.

$$C_k = \frac{1}{n-k} \sum_{i=1}^{n-k} (Q_i - \bar{Q}_i)(Q_{i+k} - \bar{Q}_i) \dots\dots (25)$$

IV. 신경망모델의 적용 및 분석

1. 대상유역 및 분석자료

본 연구에서는 일유출량해석에 대한 인공신경망모델을 적용하기 위하여 국제수문개발계획(IHP)대표유역<sup>1)</sup>중의 하나인 낙동강수계의 위천유역을 선정하여 실제수문계에 적용하였다. 위천유역의 유역면적은 472.53km<sup>2</sup>이며,

유역내에는 강우관측소 11개소와 수위관측소 6개소가 설치되어 다른 지점에 비해 비교적 자료의 습득이 용이하고 결측자료의 보완이 잘 이루어진 곳이다.

따라서 본 연구에서는 1990년이후의 1991년, 1992년, 1994년, 1995년의 자료중에서 일강우는 티센가중치에 의한 무성지점(No.1)에서 11개 우량관측소의 유역평균강우량을 사용하였고, 무성지점(No.1)의 수위자료를 수위-유량관계곡선을 이용하여 유량으로 환산하였다. 또 여기서 구한 강우, 유출량자료중에서 패턴이 비교적 잘 나타나는 6월부터 9월까지의 자료를 적용하였다.

2. 모델의 적용 및 결과

본 연구에서는 중소하천유역의 일유출량의 추계학적 비선형해석을 위하여 신경망모델을 입력층, 은닉층, 출력층으로 구성하였으며, 분석에 있어서는 1990~1995년의 일강우자료와 일유출량자료를 입력자료로 이용하였다. 그중에서 1991~1994년 자료는 학습에 의한 최적의 예측유출량을 유도하였으며, 1995년 자료는 모델의 검증분석으로 이용하였다. 또한 결과치의 통계적분석에 의하여 Table 2와 같이 나타낼 수 있다. 여기서는 일유출량의 신경망모델인 Method 1, Method 2, Method 3에 의한 결과치를 분석하여 MSE, MAPE, TSE, MAD, VER, 상대오차의 평균(Bias), 상관계수(Cor.Co.), 결정계수(Det.Co.)를 구하였다. 그 결과 대부분의 예측자료는 관측자료의 통계적 특성치를 잘 반영한 것으로 분석되었다.

본 연구에서는 1991~1994년의 학습자료에 의한 전반적인 통계분석의 결과로서 Method 3의 결정계수의 값이 0.945~0.994로서 Method 1의 0.717~0.931, Method 2의 0.777~0.938보다 우수한 결과를 나타내고 있어서, 전반적으로 모델의 효용도가 양호하다고 할 수



있다. 또한 시계열오차, 평균제곱오차, 체적오차의 백분율의 분석에서도 각각 Method 3의 경우가 -21.480~16.383%, 0.649~6.978CMS, -21.398~16.802%로 나타나서, Method 1의 24.061~45.000%, 4.297~13.269CMS, 24.003~45.002%와 Method 2의 21.004~43.200%, 2.005~14.029CMS, 20.998~43.209%보다 양호한 것으로 분석되었다. 그리고 1995년의 검증자료에 의한 통계분석결과는 Method 3의 결정계수의 값이 0.830으로서 Method 1의 0.729, Method 2의 0.659보다 우수한 결과를 나타내고 있다. 또한 시계열오

차, 평균제곱오차, 체적오차의 백분율의 분석에서도 각각 Method 3의 경우가 -31.361%, 25.885CMS, -32.167%로 나타나서, Method 1의 48.286%, 37.784CMS, 47.043%와 Method 2의 52.128%, 40.999CMS, 51.036%보다 양호한 것으로 분석되었다. 이것은 Method 3의 입력층 처리소자의 수가 Method 1, Method 2보다 많아 자료의 특성을 잘 반영하기 때문이라고 사료된다. Fig. 5는 분석자료기간 동안에 Method 1, Method 2, Method 3에 의하여 계산된 결과를 실측유출량과 함께 비교하여 도시한 것이다.

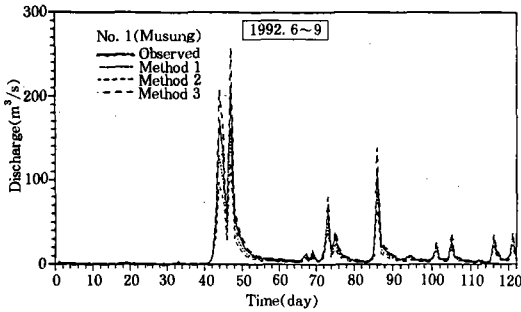


Fig. 5(a). Comparison of observed and predicted discharge at musung station (Jun.~Sep., 1992)

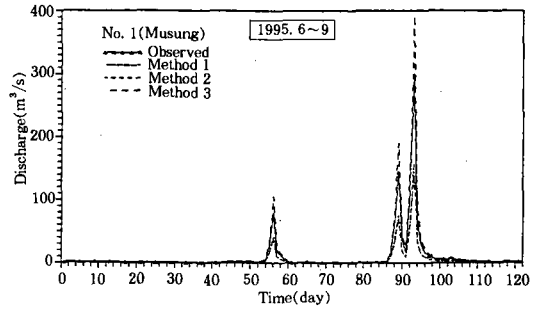


Fig. 5(b) Comparison of observed and predicted discharge at musung station (Jun.~Sep., 1995)

Table 2. The Comparison of statistical analysis between the observed and predicted discharge

Model		Index	MSE	MAPE	TSE	MAD	VER	Bias	Cor.Co.	Det.Co.
			(m <sup>3</sup> /s)	(%)	(%)	(CMS)	(%)	(m <sup>3</sup> /s)		
Method 1	Training data	1991.6~9	13.269	33.823	33.823	6.308	35.003	1.201	0.917	0.842
		1992.6~9	7.827	24.061	24.061	3.162	24.003	0.709	0.965	0.931
	Checking data	1994.6~9	4.297	45.000	45.000	2.296	45.002	0.846	0.846	0.717
Method 2	Training data	1995.6~9	37.784	48.286	48.286	6.459	47.043	3.421	0.854	0.729
		1991.6~9	10.237	26.181	26.181	4.866	27.003	0.952	0.952	0.906
	Checking data	1992.6~9	14.089	43.200	43.200	5.692	43.209	0.881	0.881	0.777
Method 3	Training data	1994.6~9	2.005	21.004	21.004	1.072	20.998	0.969	0.969	0.938
		1995.6~9	40.999	52.128	52.128	7.007	51.036	3.712	0.812	0.659
	Checking data	1991.6~9	6.369	16.383	16.383	3.028	16.802	0.982	0.982	0.964
Method 3	Training data	1992.6~9	6.978	21.480	-21.480	2.819	-21.398	0.972	0.972	0.945
		1994.6~9	0.649	6.820	6.820	0.347	6.799	0.997	0.997	0.994
	Checking data	1995.6~9	25.885	31.361	-31.361	4.416	-32.167	2.343	0.911	0.830

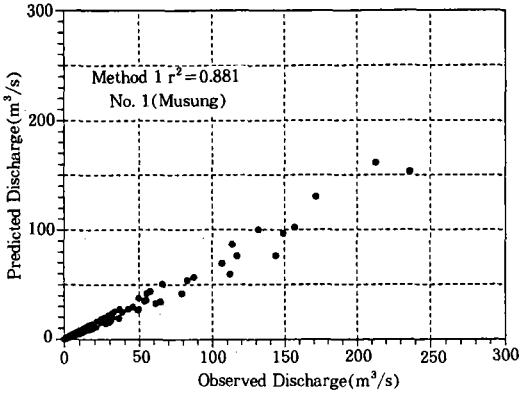


Fig. 6(a). Observed and predicted discharge using Method 1(No.1, Wi-stream)

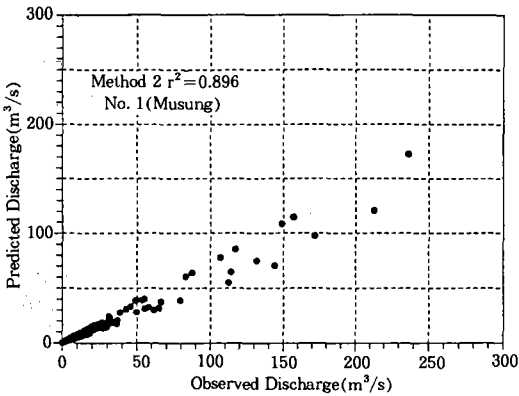


Fig. 6(b). Observed and predicted discharge using Method 2(No.1, Wi-stream)

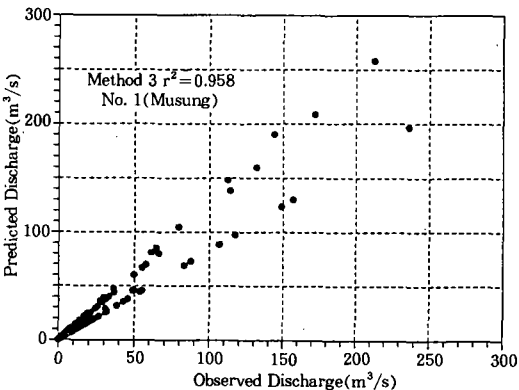


Fig. 6(c). Observed and predicted discharge using Method 3(No.1, Wi-stream)

Table 3. Comparison of frequency between daily observed and predicted discharge for each method

Class	Discharge(m <sup>3</sup> /s)		Frequency					
	Lower	Upper	Method 1		Method 2		Method 3	
	Limit	Limit	Observation	Prediction	Observation	Prediction	Observation	Prediction
1	0.00	15.00	407	437	407	438	407	408
2	15.00	30.00	40	27	40	20	40	38
3	30.00	45.00	12	9	12	16	12	15
4	45.00	60.00	11	4	11	1	11	5
5	60.00	75.00	3	1	3	5	3	6
6	75.00	90.00	3	3	3	2	3	4
7	90.00	105.00	0	3	0	1	0	2
8	105.00	120.00	4	0	4	2	4	0
9	120.00	135.00	1	1	1	1	1	2
10	135.00	150.00	1	0	2	1	2	2
11	150.00	165.00	1	3	1	0	1	1
12	165.00	180.00	1	0	1	1	1	0
13	180.00	195.00	0	0	0	0	0	1
14	195.00	210.00	0	0	0	0	0	1
15	210.00	225.00	1	0	1	0	1	1
16	225.00	240.00	0	0	0	0	0	0
17	240.00	255.00	1	0	1	0	1	0
18	255.00	270.00	0	0	0	0	0	1
19	270.00	285.00	0	0	0	0	0	0
20	285.00	300.00	1	0	1	0	1	1
Sum			488	488	488	488	488	488

본 연구에서 분석한 예측유출량과 실측유출량 사이의 상관관계를 도시하면 Fig. 6(a)~(c)와 같다. 여기에서 두 자료집단의 대응관계는 Table 2에서 분석한 상관계수 및 결정계수의 값으로 나타낼 수 있다. Table 3은 각 모델에 대한 예측유출량과 실측유출량을 함께 빈도분석한 결과이다. 이것은 예측유출량과 실측유출량의 분포특성을 알아보기 위한 것으로서, 예측유출량의 경험적 누가밀도함수(cumulative density function)와 실측유출량의 누가밀도함수를 비교하여 분석하고 이 관계를 도시하면 Fig. 7(a)~(c)와 같다. 이로부터 두 자료의 분포특성이 잘 일치하고 있음을 알 수

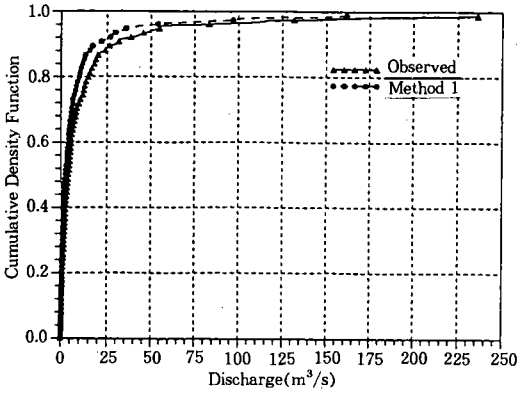


Fig. 7(a). Comparison of cumulative density function of observed and predicted discharge using Method 1

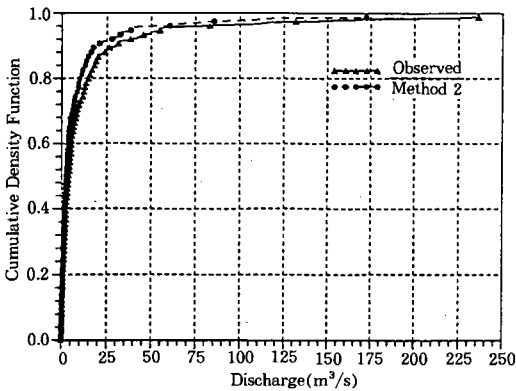


Fig. 7(b). Comparison of cumulative density function of observed and predicted discharge using Method 2

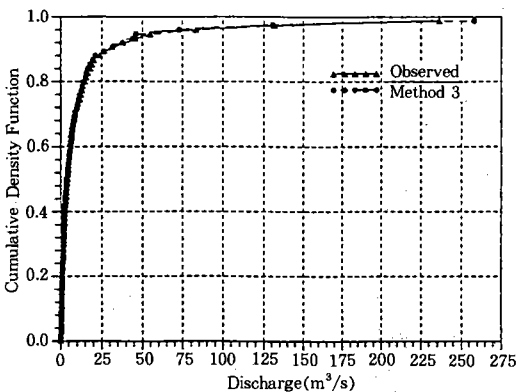


Fig. 7(c). Comparison of cumulative density function of observed and predicted discharge using Method 3

있다.

또한 수문자료의 시계열특성으로 수문학적 지속성은 1차 계열상관계수와 계열상관도를 대상으로 나타낼 수 있다. 따라서 본 연구에서 분석된 실측자료 및 예측자료의 수문시계열의 1차 계열상관계수 및 신뢰구간의 분석으로 수문학적 지속성을 나타낸 것은 Table 4와 같다. 본 연구에서는 95%신뢰구간에서의 1차 계열상관계수의 상한치와 하한치가 각각 0.168, -0.185으로 분석되었으며, 학습에 의한 분석자료인 1991년, 1992년, 1994년의 실측치, Method 1, Method 2, Method 3에 의한 1차 계열상관계수는 각각 0.4221, 0.5199, 0.7654으로 분석되었으며, 모델의 검증자료인 1995년의 실측치, Method 1, Method 2, Method 3에 의한 1차 계열상관계수는 각각 0.5124, 0.5134, 0.5133, 0.5138으로 분석되어서 신뢰구간의 하한계와 상한계의 범주에 속하지 않으므로 실측 및 예측유출량은 수문학적 지속성을 가진다고 할 수 있다. 또한 계열상관도에 의한 예측유출량과 실측유출량의 계열상관특성을 비교한 것은 Fig. 8(a)~(c)와 같고, 각 모델들의 예측유출량의 계열상관계수가 실측유출량의 계열상관계수와 거의 일치하여 예측유출량이 실측유출량의 수문학적 지속성을 잘 재현하고 있음을 알 수 있었다.

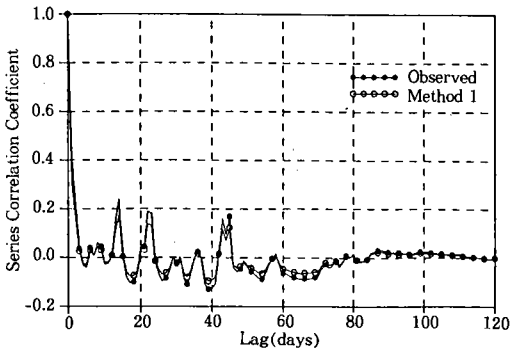
## V. 결 론

본 연구에서는 하천유역의 비선형 추계학적 일유출해석에 인간의 사고과정을 수학적으로 모델화한 신경망모형이론을 수문학적으로 도입하여서 적용하고 이를 검토하였다. 분석에 있어서는 국제수문개발계획(IHP) 대표시험유역인 위천유역의 일강우량 및 일유출량자료를 이용하여 학습자료 및 검증자료로 구분하여 이용하였으며, 실제 하천유역에서 적용할 수 있는 일유출량예측 신경망모델을 구성하여,

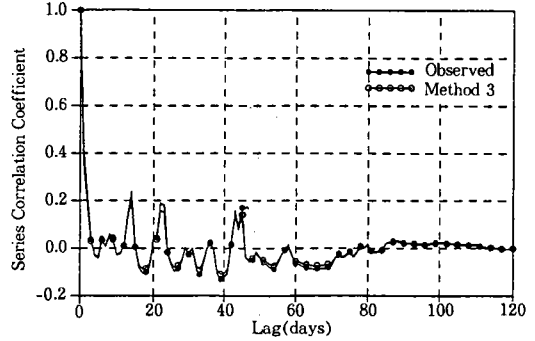
**Table 4. Comparison of lag-one serial correlation coef. between observed and predicted discharge**

Case		Lag-one serial correlation coefficient( $r_1$ )								Remark
		Observed		Method 1		Method 2		Method 3		
Training data	1991	0.4221	○	0.4221	○	0.4221	○	0.4221	○	Upper lag-one serial C.C : 0.168 Lower lag-one serial C.C : -0.185
	1992	0.5199	○	0.5199	○	0.5199	○	0.5199	○	
	1994	0.7654	○	0.7654	○	0.7654	○	0.7654	○	
Checking data	1995	0.5124	○	0.5138	○	0.5133	○	0.5138	○	

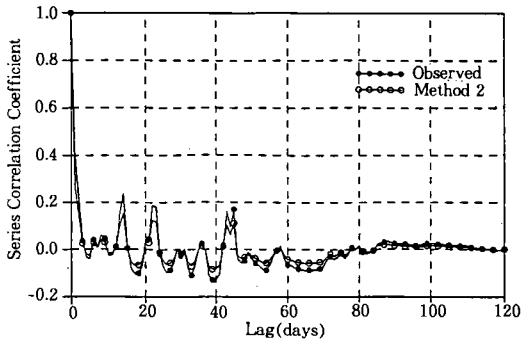
Note) ○ : 수문학적 지속성을 가짐, × : 수문학적 지속성을 가지지 않음,  
Lag-one Serial C. C. : 1차 계열상관계수( $r_1$ )



**Fig. 8(a). Comparison of series correlation coefficient of observed and predicted discharge using Method 1**



**Fig. 8(c). Comparison of series correlation coefficient of observed and predicted discharge using Method 3**



**Fig. 8(b). Comparison of series correlation coefficient of observed and predicted discharge using Method 2**

다음과 같은 결론을 유도하였다.

1) 하천유역의 비선형 추계학적 일유출량해석을 위하여 신경망모델을 입력층, 은닉층 및 출력층으로 구성하고, 신경망모델의 학습에

의한 최적의 일유출량을 유도하였다. 또한 모델의 검증을 통하여 입력층 처리소자의 수가 많을수록 관측치의 특성을 잘 나타내며, 그 결과가 양호한 것으로 나타났다.

2) 신경망모델에 의해 분석된 예측유출량과 실측유출량의 수문학적 지속성을 판정하는 기준으로 1차 계열상관계수 및 계열상관도에 의한 분석을 실시하였으며, 그 결과 수문학적 지속성을 가지는 유량자료계열로 나타났다. 특히 계열상관도에 의한 유량자료계열의 계열상관계수는 거의 일치하는 것으로 분석되었다.

3) 신경망이론을 적용한 일유출량 모델에 의해 분석된 예측유출량과 실측유출량을 비교·분석하기 위하여 시계열오차, 평균제곱오차, 체적오차의 백분율 및 결정계수등의 통계적분석의 결과로서 신경망모델에 대한 적합성이

뛰어난 것으로 분석되었고, 특히 Method 3의 신경망모델은 실측치에 가장 근접한 모델로 분석되었다.

4) 본 연구에서는 하천유역의 물리적 복잡성에 의하여 수학적 모형의 적용이 어려운 수문현상을 신경망이론을 도입하여 일유출량을 분석한 결과, 모델의 구성 및 운영이 간편하고 적용성이 뛰어난 것으로 분석되었다. 따라서 이를 기본으로 하여 실시간을 고려한 홍수 예측을 위한 적응 네트워크 알고리즘의 개발 및 담운영을 위한 최적의 의사결정 시스템에 관련된 적용분야에 대하여 지속적인 연구가 필요한 것으로 사료된다.

### 참 고 문 헌

1. 건설교통부, 1988~1996, 국제수문개발계획(IHP) 대표유역 연구조사보고서.
2. 김주환, 1993, 신경회로망을 이용한 하천 유출량의 수문학적 예측에 관한 연구, 인하대학교 박사학위논문.
3. 오남선, 선우중호, 1996, 신경망이론에 의한 강우예측에 관한 연구, 한국수자원학회지, 제29권 제4호, pp. 109~118.
4. 이남호, 정하우, 1994, 신경회로망을 이용한 수도증발산량 예측-백 프로파게이션과 카운터 프라파게이션 알고리즘의 적용, 한국농공학회지, 제36권 제2호, pp. 88~95.
5. Anderson, R. L., 1941, Distribution of the serial correlation coefficients, Annual of Math. Statistics, Vol. 8, No. 1, pp. 1~13, March.
6. Hertz, J., Krogh, A., and Plamer, R. G., 1991, Introduction to the theory of Neural Computation, Addison-Wesley Pub. Co.
7. Jason Smith and Robert N. Eli, 1995, Neural-Network Models of Rainfall-Runoff Process, J. of Water Resour. Plng. and Mgmt., ASCE, Vol. 121, No. 6, pp. 499~508.
8. Karunanithi, N., Grenney, W. J., Whitley, D., and Bovee, K., 1994, Neural networks for river prediction, J. of Comp. in Civil Engineering, ASCE, Vol. 8, No. 2, pp. 201~220.
9. Lapedes A.S. and Farber R., 1987, Non-Linear Signal Processing using Neural Networks : Prediction and System Modeling, Los Alamos National Laboratory Report, LA-UR-87-2662.
10. Smith, J., 1992. Streamflow forecasting using a backpropagation neural network, M. S. thesis, West Virginia Univ., Morgantown, W. Va.
11. Yoh-Han Pao, 1989, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Publishing Co., Inc., pp. 269~290.