

〈論 文〉

패턴 認識 方法을 적용한 河川流出의 非線形 豫測

Nonlinear Prediction of Streamflow by Applying Pattern Recognition Method

강 관 원*, 박 찬 영**, 김 주 환***

KANG Kwan Won, PARK Chan Yung and KIM Ju Hawn

Abstract □ The purpose of this paper is to introduce and to apply the artificial neural network theory to real hydrologic system for forecasting daily streamflows during flood periods.

The hydrologic dynamic process of rainfall-runoff is identified by the iterated estimation of system parameters that are determined by adjusting the weights of the network according to the non-linear response characteristics which is formed the model.

Back propagation algorithm of neural network model is applied for the estimation of system parameters with past daily rainfall and runoff series data, and streamflows are forecasted using the parameters.

The forecasted results are analyzed by statistical methods for the comparison with the observed.

要 旨 : 본 연구는 홍수기의 일단위 하천유출량을 예측하기 위한 방법으로 인공지능의 구현 모형으로 사용되고 있는 신경회로망이론을 도입하여 實水文系에 적용하고 그 결과를 제시하는 것이다.

강우-유출과정으로 형성되는 수문계의 동적거동을 입력패턴으로 보아서 모형을 구성하는 유니트의 비선형 응답특성에 따라 네트워크의 상호 결합강도를 조정하여 시스템의 매개변수를 반복추정하는 방법으로 시스템을 等定 평가하였다.

일강우와 일유량의 과거 관측치를 신경회로망 모형의 逆傳播알고리즘으로 학습시켜 추정된 매개변수를 이용하여 하천유출량을 예측하였고 그 결과를 관측된 유량과 비교하기 위하여 통계학적으로 분석하였다.

1. 서 론

이수 및 치수를 목적으로 하는 수자원 관리 및 운영을 위한 계획에서 발생하는 문제의 해결을 위한 기본적인이고 중요한 요소중의 하나는 하천유출량의 정확한 추정이라 할 수 있다. 하천유출량은 주로 강우가 하천유역이라는 대단히 복잡한 시스템을 통하여 발생하는 결과로 볼 수 있다. 하천유역은 여러가지 수문기상학적인자를 포함하는 시스템으로 간주할 수 있으며 여기에서 발생하는 강우-유출현상은 비정규성과 불확실성을 내포하기 때문에 강한 비선형성을 나타낸다.

예측이란 시스템의 동적상태를 수학적으로 표현하는

시스템등정에 근거하여 미래의 상태를 알아내는 것이다. 따라서 정확한 하천유출량의 예측은 얼마나 정확히 시스템을 등정하였는가에 달려있다. 수학적 모형으로 개발된 신경회로망은 원래의 목적인 패턴인식뿐만 아니라 그 함수가 가지고 있는 비선형 학습 및 보간능력으로 인하여 응용분야에서 광범위하게 사용되고 있다. 신경회로망은 시스템의 상태와 매개변수들의 최적한 값들을 반복적으로 순환추정할 수 있는 반복순환 알고리즘을 가지며 시스템을 구성하는 비선형 특성함수로 인하여 비선형시스템의 모형화에 이용할 수 있다.

본 연구에서는 강우-유출과정을 유역이라는 시스템

*인하대학교 토목공학과 교수

**인하공업전문대학 토목과 교수

***인하대학교 대학원 토목공학과 박사과정 수료

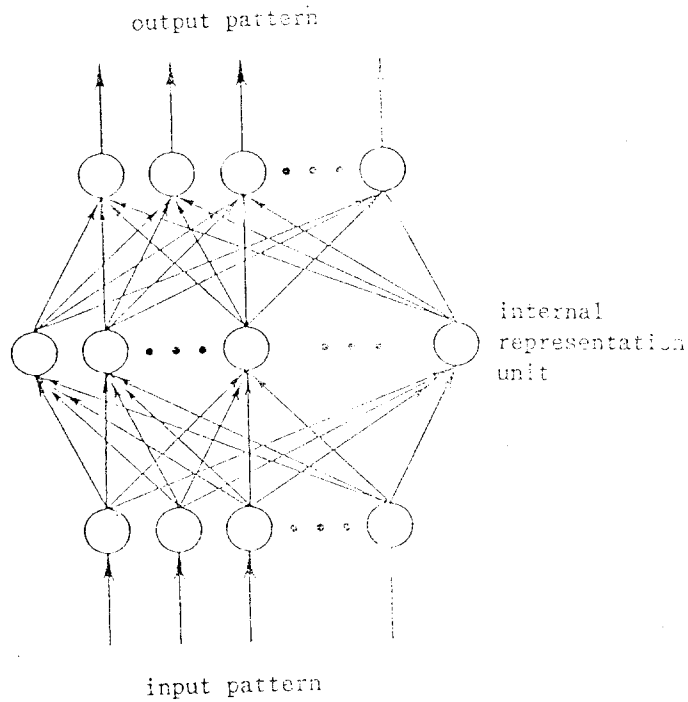


Fig.1 Network Architecture

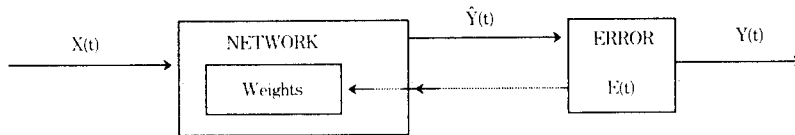


Fig.2 Block diagram of error back propagation algorithm

내에서 발생된 신호체계의 패턴으로 간주하고 모형을 구성하는 각 유니트간의 매개변수를 인공신경회로망 모형의 역전파알고리즘으로 최적화하는 방법을 實際水文系에 적용하고 홍수기의 일단위 하천유출량의 과거기록치를 이용하여 유출량을 예측하고자한다. 또 그 결과를 비교분석하여 수문량 예측에 대한 적용가능성을 제시하고자 한다.

2. 적용 예측이론

2.1 패턴인식 방법에 있어서의 인공신경회로망

그물과 같이 연결된 뇌의 정보처리과정의 동적거동

을 표현하기 위하여 개발된 인공신경회로망은 여러 정보, 통신, 제어 분야의 복잡한 문제에 적용되고 있다. 인간의 두뇌를 형성하고 있는 신경조직은 뉴런(neuron)이라는 신경소자가 복잡하게 얽힌 비선형 시스템으로 이들을 서로 연결해 주는 결합강도에 의하여 시스템의 성능을 향상 시킨다. 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 표현될 수 있으며 입력과 출력간의 상관관계는 뉴런간의 결합강도에 의존한다. 결합강도는 학습과 자기조직에 의하여 조정(adjustment)된다. 신경회로망은 이러한 뇌와 동일한 계

산방식을 사용한 커넥셔니즘(connectionism)에 기반을 둔 數理模型이다.

2.2 퍼셉트론에 의한 예측 학습알고리즘

퍼셉트론(perceptron)은 미국의 Rosenblatt가 1957년 패턴을 분류하기 위하여 제안했던 신경회로망 모형의 패턴인식 방법으로 그림 1과 같이 입력층, 중간층, 출력층으로 구성된 계층적 구조를 갖는다. 다층퍼셉트론(multi-layer perceptron)은 1986년 Rumelhart에 의하여 제안되었으며 중간층에 어느 정도 뉴런을 비교적 많이 둔 것으로 각 패턴마다 중간층에 필요한 정보를 학습시킬 수 있도록 한 것이다.

입력층은 대상 패턴에 대하여 각 입력값에 결합강도를 곱해서 중간층으로 전달한다. 중간층은 입력층으로부터 받은 값들을 모두 합하여 전달함수를 통하여 출력하게 되고 이 값들은 다시 중간층과 출력층사이의 결합강도와 곱하여져 출력층으로 보내진다. 逆傳播 알고리즘(back propagation algorithm)은 다층 퍼셉트론에서 입력패턴과 출력패턴에 대하여 오차를 줄여가는 과정에 적용되는 방법으로 이는 정보처리 시스템의 목적에 맞도록 합리적으로 각 층에서 입출력 패턴의 결합강도를 조절하는 것이다. 일반화된 델타룰(generalized Delta rule)이라고도 불리우는 이 과정에서는 출력층에서 출력되는 출력값과 실제값의 자승오차가 감소되는 방향으로 피이드백 되며 결합강도가 차츰 최적치에 가까워질 때까지 반복되어 추정된다.

여기에서 출력의 결합강도의 변화량은 출력의 자승오차의 합을 최소화하기 위하여 오차가 감소하는 방향으로 구성된다.

입력층 I에 패턴 p가 입력되었을 때 출력층 O로 출력되는 네트워크가 있다고 하자. 여기서 출력층의 유닛 k의 출력 net_{pk} 는 I층 유닛 i의 출력 I_{pi} 와의 결합강도 W_{ki} 와의 결합강도의 합에 의해 결정된다. 출력층의 sigmoid를 거치지 않은 값을 net_{pk} 라 하고 거친 값을 O_{pk} 라 한다면 다음 식과 같이 쓸 수 있다.

$$net_{pk} = \sum_i W_{ki} I_{pi} \tag{1}$$

$$O_{pk} = sig(net_{pk}) \tag{2}$$

여기서, p : 패턴 번호

k : 출력층 유닛

i : 입력층 유닛

출력층에서 출력되는 값과 실제값과의 자승오차의 합이 오차함수가 되며 패턴 p에 대한 오차함수 E_p 와 모든 패턴에서의 오차 E_k 는 다음 (3), (4)식과 같이 쓸 수 있다.

$$E_p = \frac{1}{2} \sum_k (T_{pk} - O_{pk})^2 \tag{3}$$

$$E_k = \sum_p E_p \tag{4}$$

여기서 T_{pk} 는 p번째 입력패턴의 k번째 유닛의 실제값이며 오차 E_k 가 최소로 되는 상태를 최적 네트워크라고 오차 E_p 를 결합강도 W_{ki} 에 관하여 최소화하는 것으로 한다. 오차 E_p 는 O층의 출력 O_{pk} 의 함수이고 그 출력 O_{pk} 는 I층의 출력 I_{pi} 의 함수이다. 중간층과 출력층에서는 출력의 자승오차의 합을 최소화하기 위하여 결합강도의 변화량은 관계식 (5)와 같이 오차가 감소하는 방향으로 형성된다.

$$\Delta W_{kj} \propto - \frac{\partial E_p}{\partial W_{kj}} \tag{5}$$

j : 중간층 유닛

위 식에서 최솟치가 얻어질 수 있는 방법으로 最急降下法(gradient descent method)을 적용하여 (6)식으로 표현되는 결합강도의 보정량을 구한다.

$$\Delta W_{kj} = \eta \delta_{pk} O_{pj} \tag{6}$$

여기서, δ_{pk} : p번째 입력패턴의 k번째 유닛의 델타값
 η : 학습율

식(6)에서 δ_{pk} 는 다음 (7)식으로 정리된다.

$$\delta_{pk} = (T_{pk} - O_{pk}) * O_{pk} * (1 - O_{pk}) \tag{7}$$

다층 신경회로망의 경우 중간층의 에러를 결정하여 주어야하며 이를 위하여 출력층의 에러를 반영하여 중간층의 δ_{pj} 를 결정한다.

$$\delta_{pj} = O_{pj} * (1 - O_{pj}) \sum_k \delta_{pk} W_{kj} \tag{8}$$

학습 과정은 각 입력값에 대해 출력을 계산하는 순방향처리와 실제값과 출력값 사이에서 발생하는 오차

를 역으로 진행시켜 이 오차를 최소화하도록 층과 층 사이의 결합강도를 변화, 조절하는 역방향처리로 구성된다. 결합강도 W_{kj} , W_{ji} 의 초기치는 -0.1과 0.1사이의 난수를 사용한다.

순방향처리는 입력자료가 신경회로망을 통하여 출력 자료를 계산하는 과정이 입력층, 중간층, 출력층을 차례로 통과하여 출력값을 산출하는 과정이다.

역방향처리과정은 출력층에서의 실제값과 출력값과의 오차를 최소화하도록 결합강도 W_{kj} 를 조절하고 중간층의 출력값을 얻어 중간층과 입력층 사이의 결합강도 W_{ji} 를 반복적으로 조절하는 형태를 갖는다.

출력층과 중간층 사이의 결합강도의 조절은 아래 (9)식과 같다.

$$\Delta W_{kj}(n) = \eta * \delta_{pk} * O_{pj} + \alpha * \Delta W_{kj}(n-1) \quad (9)$$

여기서, n : 반복회수

α : 모멘텀상수

따라서 다음 반복 계산을 위하여 조절된 결합강도는 다음과 같다.

$$W_{kj}(n+1) = W_{kj}(n) + \Delta W_{kj}(n) \quad (10)$$

중간층과 입력층 사이의 결합강도의 변화량은 (11)식과 같고 다음 계산을 위하여 조절된 결합강도는 (12)식과 같다.

$$\Delta W_{ji}(n) = \eta * \delta_{mj} * O_{mi} + \alpha * \Delta W_{ji}(n-1) \quad (11)$$

$$W_{ji}(n+1) = W_{ji}(n) + \Delta W_{ji}(n) \quad (12)$$

출력층의 총오차 ($E = \sum E_p$)가 임의의 오차허용치를 만족하게되면 반복학습을 끝낸다.

3. 적용자료

본 연구에서는 IHP대표유역인 평창강유역의 11개 강우관측기록자료(1985-87)를 평균강우로 환산하였고 방림지점의 수위자료(1985-87)를 수위-유량관계곡선을 이용하여 유량으로 환산하였다. 또 여기서 얻은 강우, 유량자료중에서 패턴이 비교적 잘 나타나는 6월부터 9월까지의 자료를 적용하였다.

방림지점의 수위-유량 관계 곡선(IHP 1990)은 다음

과 같다.

$$Q = 106.057(h - 0.647) \quad h > 0.7m$$

$$Q = 0.426h \quad h < 0.7m$$

이를 적용하기 위하여 대상자료의 자기상관함수와 교차상관함수를 구하여 통계학적으로 분석하므로써 자료 상호간의 상관성을 고려하여 모형의 입력, 출력패턴을 결정하였다.

4. 모형의 구성

모형은 자료의 상관성을 분석하여 다음과 같이 서로 다른 몇가지 형태로 표현하였다.

MODEL I: $y_{k+1} = f(y_k, u_{k-1}, u_k; \text{Boolean input})$

MODEL II: $y_{k+1} = f(y_{k-1}, u_k, u_{k-1}, u_k; \text{Boolean input})$

MODEL III: $y_{k+1} = f(y_{k-2}, y_{k-1}, y_k, u_{k-1}, u_k; \text{Boolean input})$

MODEL IV: $y_{k+1} = f(y_{k-2}, y_{k-1}, y_k, u_{k-1}, u_k, u_{k+1}; \text{Boolean input})$

여기서, y : 유량

u : 강우량

Boolean input : 월별 구분을 위한 index input

여기서 Boolean 입력은 사용된 자료가 6월~9월까지이므로 년도가 바뀔 때마다 월별 구분을 위하여 각 월에 대하여 4가지로 구분하여 입력하였고 일련의 연속된 자료기간이 아니라는 것을 인식하도록 하기위한 것이다. 따라서 신경회로망에 적용될 모형을 구성하는 네트워크는 이들 4개의 패턴이 포함되며 각 변수에 대하여 표 1과 같이 구성하였다.

5. 적용결과 및 고찰

표 1로 구성되는 모형에 의하여 계산된 결과((a) 관측치와 예측치와의 대응관계, (b) 자기상관함수, (c) 각 년도별 실측치와 예측치의 비교)는 그림 3~4에서 각 년도별로 도시하였다. 그림 3은 모형 III에 의하여 계산된 결과이며 그림 4는 모형 IV에 의하여 계산된 결과를 나타낸 것이다. 또 예측치가 실측치를 얼마나 잘 재현해 주는가를 알아보기 위하여 빈도분석을 한 결과

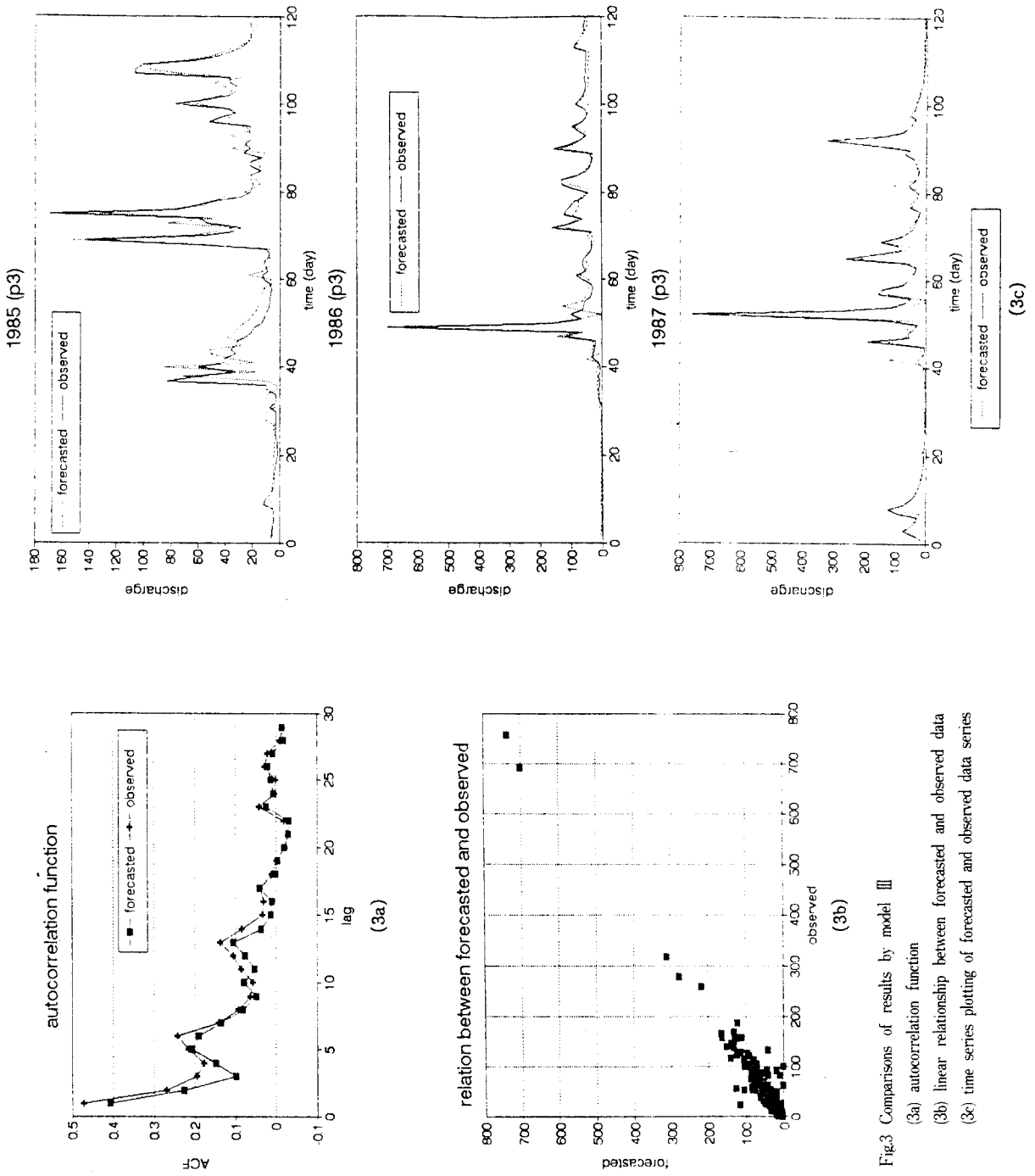


Fig.3 Comparisons of results by model III
 (3a) autocorrelation function
 (3b) linear relationship between forecasted and observed data
 (3c) time series plotting of forecasted and observed data series

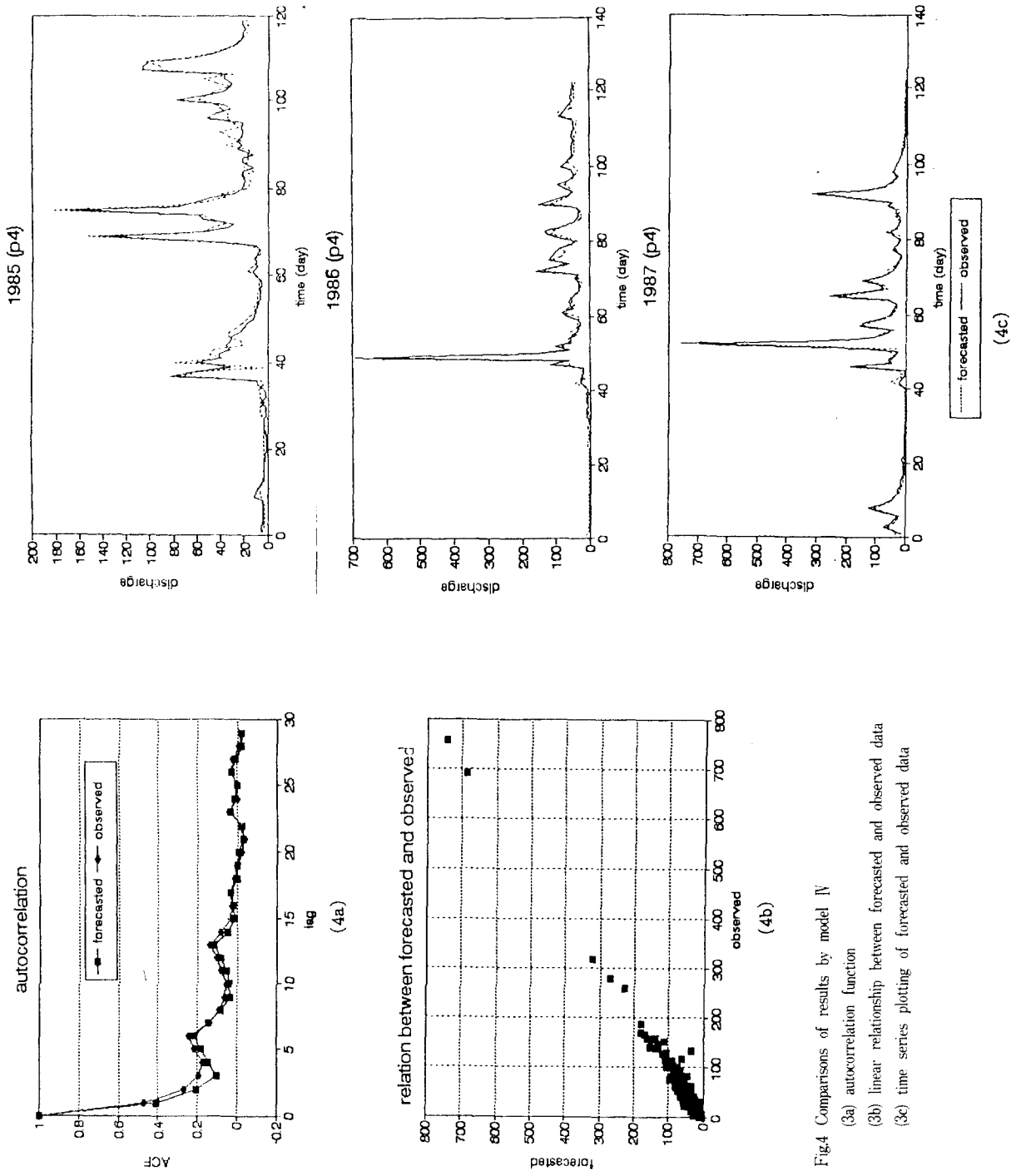


Fig.4 Comparisons of results by model IV
 (3a) autocorrelation function
 (3b) linear relationship between forecasted and observed data
 (3c) time series plotting of forecasted and observed data

Table 1 Formation of network for each model

구 분	input unit數	hidden unit數	output unit數	비 고
Model I	7	14	1	
Model II	8	16	1	
Model III	9	18	1	
Model IV	10	20	1	

Table 2 Frequency table of the observed and forecasted data set

Class	Lower Limit	Upper Limit	Frequency				
			OBS.	Model I	Model II	Model III	Model IV
1	.00	29.23	194	199	201	194	195
2	29.23	58.46	106	92	105	101	94
3	58.46	87.69	35	38	30	38	36
4	87.69	116.92	12	18	12	11	15
5	116.92	146.15	6	8	7	11	10
6	146.15	175.38	3	4	4	3	7
7	175.38	204.62	2	1	1	0	1
8	204.62	233.85	1	1	0	1	0
9	233.85	263.08	0	0	0	0	1
10	263.08	292.31	1	1	1	1	1
11	292.31	321.54	1	1	1	1	1
12	321.54	350.77	0	1	0	0	0
13	350.77	380.00	0	0	0	0	0
14	380.00	409.23	0	0	0	0	0
15	409.23	438.46	0	1	0	0	0
16	438.46	467.69	0	0	0	0	0
17	467.69	496.92	0	0	0	0	0
18	496.92	526.15	0	0	0	0	0
19	526.15	555.38	0	0	0	0	0
20	555.38	584.62	0	0	0	0	0
21	584.62	613.85	0	0	0	0	0
22	613.85	643.08	0	0	0	0	0
23	643.08	672.31	0	0	0	0	0
24	672.31	701.54	1	1	1	1	1
25	701.54	730.77	0	0	0	0	0
26	730.77	760.00	1	0	1	1	1

는 표 2와 같다.

표 3은 각 모형에 의하여 계산된 자료의 통계적 특성을 수록한 것이다. 여기서는 모형의 구조를 표현하는 패턴의 수가 많을수록 실측치와 비교하여 볼 때 유

사한 것으로 나타났다.

그림 5는 (4)식으로 표현되는 시스템의 오차를 도시한 것이다. 여기서 시스템의 오차는 매개변수인 결합강도를 추정할 때 반복회수에 따라 계속 감소하며 증가하지 않음을 알 수 있다. 이는 Rosenblatt의 수렴정리를 증명해 주며 어떠한 시스템에서도 오차를 줄여 나갈 수 있음을 의미한다.

6. 결 론

본 연구에서는 인공지능모형의 구현을 위하여 사용되고 있는 패턴인식 알고리즘을 강우-유출과정에 적용하여 그 결과를 제시하였다. 그 결과 강우-유출과정을 시스템 등정할 수 있는 가능성이 있다는 것을 알 수 있었다.

입출력시스템의 구성에서 포함되는 비선형 전달함수로 인하여 강우-유출과정이 가지는 비선형성을 고려할 수 있고 유출량 예측에 있어 입력유니트의 과거 패턴에 가장 유사한 경우에 근거하여 예측치를 출력하므로 하천유역의 강우-유출모형화 및 예측에 적용할 수 있을 것으로 판단된다.

인공 신경회로망시스템은 시스템의 매개변수 추정방법으로서 학습에 의하여 시스템의 국부적 변형 및 잡음에 민감하지 않으므로 강우, 유출 자료와 같이 잡음을 포함하며 왜곡된 그리고 애매한 자료를 효과적으로 처리할 수 있고 주어진 자료에 대한 적합능력이 뛰어난 것을 알 수 있었다. 따라서 자연현상과 같은 수학적 알고리즘의 적용이 곤란한 문제를 효과적으로 해결할 수 있는 가능성이 있다.

모형의 입출력 패턴을 구성하여 계산한 결과 강우-유출시스템을 표현하는 입출력 자료의 패턴 수가 많을수록 관측치의 특성을 잘 표현해 주며 그 결과가 양호한 것으로 나타났다.

따라서 이를 기본 이론으로 하여 실시간을 고려한 홍수예측을 위한 적응 네트워크 알고리즘의 개발, 그리고 예측된 홍수량에 대하여 댐 운영을 위한 최적의 의사결정 시스템에 관련된 적용분야에 대한 연구는 앞으로의 과제이다.

감사의 글

Table 3 Comparison of statistical properties

구	분	Mean	Median	STD	Min	Max	상관계수	결정계수(R ²)
Model I	OBS.	40.63	25.78	66.81	0.08	757.77	0.946	0.896
	FOR.	39.95	25.66	56.23	0.12	696.86		
Model II	OBS.	40.72	25.78	66.88	0.08	757.77	0.970	0.941
	FOR.	37.50	24.10	63.66	0.08	755.25		
Model III	OBS.	40.81	25.78	66.95	0.08	757.77	0.971	0.941
	FOR.	39.46	26.14	64.09	0.13	736.79		
Model IV	OBS.	40.81	25.78	66.95	0.08	757.77	0.983	0.966
	FOR.	39.01	24.96	64.28	0.08	744.13		

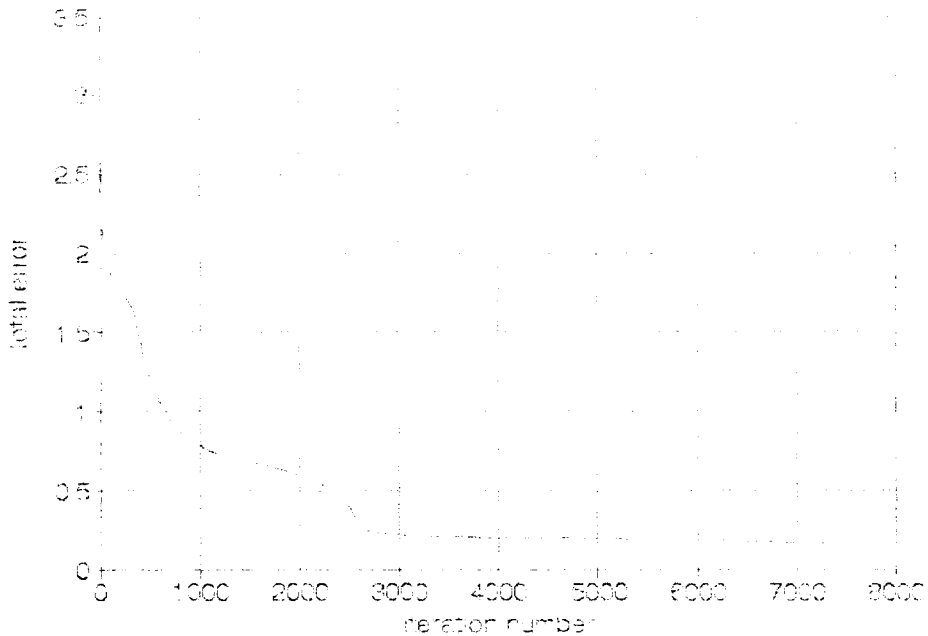


Fig. 5 Variation of total error according to iteration number

本 研究는 本校 附設 産業科學技術研究所의 支援으로 이루어졌으며 當局에 깊은 感謝를 드립니다.

參考文獻

1. Andres S. Weigend, et al., "Back-Propagation, Weight-Elimination and Time Series Prediction," SMC, Connectionist Models Proceedings of the 1990 summer school, pp.105-116, 1990.
2. Box, G.E.P and G.M. Jenkins, "Time Series Analysis, Forecasting and Control," Holdenday, San Francisco, 1970
3. Benito Fernandez, et al., "Nonlinear Dynamics System Identification using Artificial Neural Networks," IJCNN, June, Vol.2, pp. II-133-II-142, 1990.
4. Barto Kosko, "Neural Networks and Fuzzy Systems," Prentice Hall, 1992.
5. Chiu, C. L. and Huang, J. T. "Nonlinear Time Varying Model of Rainfall Runoff Relation," W. R. R., Vol.6, 1970
6. Dawdy, D. R. "The Progress of Hydrology, Mathematical Modeling in Hydrology," Proceedings of 1st Int. Seminar for Hydrology Professors, Vol. 1, pp.346-361, 1969.

7. D.Ouazar and C.A.Brebba, "Computer Methods and Water Resources: Computational Hydrology," 1st Int. Conference, Morocco Vol.3, 1988
8. David B. Parker, "Optimal algorithms for adaptive networks: second order back propagation, second order direct propagation, and second order Hebbian learning," Proceedings of the 1st International Conference on Neural Networks, San Diego, Calif., June, vol2, pp II-593-II-600, 1987.
9. Hino, M "Runoff Forecasts by Linear Predictive filter," Proc. ASCE, J. of Hydraulic Div., Vol.96, No. HY3, pp 681-701, 1970.
10. John Hertz, Andes Krogh, Richard G. Palmer, "Introduction to the Theory of Neural Computation," Addison-Wesley Publishing Company, 1991.
11. Lapedes, A. & Farber, R., "Non-Linear Signal Processing using Neural Networks: Prediction and System Modeling," Los Alamos National Laboratory report, LA-UR-87-2662.
12. Philip D. Wasserman, "Neural Computing Theory and Practice," Van Nostrand Reinhold, 1989.
13. Takashi Kimoto, Kauzo Asakawa, Morio Yoda and Masakazu Takeoda, "Stock Market Prediction System with Modular Neural Network," INNS, Vol.1, pp. I-1 - I-6, 1990.
14. Yoh-Han Pao, "Adaptive Pattern Recognition and Neural Networks," Addison-Wesley Publishing Company, pp.269-290, 1989.
15. Eric A. Wan, "Temporal Backpropagation: An Efficient Algorithm for Finite Impulse Response Neural Networks," SMC, Connectionist Models Proceedings of the 1990 summer school, pp.131-137, 1990.
16. Eric F. Wood et.al, "Real Time Forecasting/Control of Water Resource System," selected papers from IIASA Workshop, Vol.8 Oct. 18-20, 1976.
17. Marcus D. Odom, Ramesh Sharda, "A Neural Network Model For Bankruptcy Prediction," IJCNN, June, Vol.2, pp. II-163-II-168, 1990.
18. Reza Shadmehr and David Z. D'Argeno, "A Comparison of a Neural Network Based Estimator and Two Statistical Estimators in a Sparse and Noisy Data Environment," IJCNN, January, Vol.1, pp. I-289-I-292, 1990.
19. Robert Hecht-Nielsen, "Neurocomputing," Addison-Wesley Publishing Company, 1989.
20. V. P. Singh, "Rainfall Runoff Relationship," Water Resources Publications, 1982.
21. Liu, C. C. K., and W. Brutsaert, "A Nonlinear Analysis of the Relationship Between Rainfall and Runoff for Extreme Floods," W. R. R. Vol. 14, No. 1, 1978.