韓國水資源學會論文集
第41卷 第10號 · 2008年 10月
pp. 1035~1044

DOI: 10.3741/JKWRA.2008.41.10.1035

유출량 및 수질자료를 이용한 인공신경망 예측모형 개발에 관한 연구

Study on Development of Artificial Neural Network Forecasting Model Using Runoff, Water Quality Data

오 창 열* / 진 영 훈** / 김 동 렬*** / 박 성 천****

Oh, Chang Ryeol / Jin, Young Hoon / Kim, Dong Ryeol / Park, Sung Chun

Abstract

It is critical to study on data charateristics analysis and prediction for the flood disaster prevention and water quality monitoring because discharge and TOC data in a river channel are strongly nonlinear. Therefore, in the present study, prediction models for discharge, TOC, and TOC load data were developed using approximation component in the last level and detail components segregated by wavelet transform. The results show that the developed model overcame the persistence phenomenon which could be seen from previous models and improved the prediciton accuracy comparing with the previous models. It might be expected that the results from the present study can mitigate flood disaster damage and construct active alternatives to various water quality problems in the future.

keywords : Wavelet Transform, Approximation Component, Detail Components, Artificial Neural Networks, Total Organic Carbon, Persistence Phenomenon

요 지

하도내에서 발생하는 유출량 및 TOC 자료는 비선형성이 강한 자료임에 따라 홍수에 대한 재난대응과 수질의 상시감시를 위해서는 자료의 특성 분석과 예측에 관한 연구는 필수라 할 수 있다. 따라서 본 연구에서 유출량 및 TOC, TOC부하량 자료에 대한 웨이블렛 변환에 의해 최종분해된 최종과형분해단계의 근사성분과 상세성분을 이용하여 예측모형을 개발하였다. 그 결과 기존 인공신경망 모형에서 관찰되었던 시계반대 방향으로 전이되는 지속현상의 극복 가능성을 보여주었으며, 기존 인공신경망 모형에 비하여 예측의 정확도가 향상됨을 확인할 수 있었다. 이러한 연구결과는 향후 홍수에 대한 피해를 최소화하고 각종 수질사고에 적극적인 대응방안 수립이 가능할 것으로 기대된다.

핵심용어 : 웨이블렛변환, 근사성분, 상세성분, 인공신경망, 총유기탄소, 지속현상

* 유량조사사업단 품질정책실 선임연구원
Senior Researcher, Quality Control & Policy Div., Hydrological Survey Center, Goyang, Kyeonggi, Korea.
(e-mail: croh@kict.re.kr)

** 교신저자, 동신대학교 토목공학과 연구교수
Research Professor, Dept. of Civil Engineering, Dongshin University, Naju, Jeonnam, Korea,
(e-mail: nmdr@freechal.com)

*** 동신대학교 토목공학과 박사과정
Doctor Course, Dept. of Civil Engineering, Dongshin University, Naju, Jeonnam, Korea,
(e-mail: kdr86@mltm.go.kr)

**** 동신대학교 토목공학과 교수
Professor, Dept. of Civil Engineering, Dongshin University, Naju, Jeonnam, Korea,
(e-mail: psc@dsu.ac.kr)

1. 서 론

유출량은 수자원 생성의 기본요소인 강우현상에 대한 유역의 반응으로 발생하며 이에 대한 해석과 예측은 강우-유출 과정(rainfall-runoff process)에 대한 보다 깊은 이해가 이루어져야 한다. 그러나 강우-유출 과정의 연결 매체인 유역은 복잡한 특성인자들로 구성되어 있음에 따라 비선형적 변환을 유발하며 각각의 변수들 역시 시공간적으로 그 변동이 매우 심하다. 총유기탄소(Total Organic Carbon; 이하 TOC) 수질자료도 생태계의 물리적·화학적·생물학적으로 매우 복잡한 반응기작에 의해 생성된 비선형적 특성이 강한 자료이다.

이처럼 비선형적인 특성을 갖는 자료에 대한 해석은 기존 선형모형에서는 무시되어 왔으나, 최근 전자계산능력의 발달로 이에 대한 연구가 가능하게 되었다. 이러한 성과는 수문분야에서도 활발히 진행되고 있으며 이에 대한 연구 성과를 살펴보면 다음과 같다.

인공신경망 이론은 평창강 유역에 대한 유출량 예측에 관한 연구(김주환, 1993)를 시초로 하여 수자원 및 환경공학 분야에서 예측을 위한 도구로서 그 활용방법이 다양하게 이용되고 있다. 유출량 예측에 관한 연구를 유역별로 살펴보면 한강 유역(윤강훈 등, 2004c), 낙동강 유역(윤강훈 등, 2004a; 2004b), 영산강 유역(이관수 등, 2000) 및 설마천 유역(이정규 등, 2003)에 적용하여 수문학적 예측의 우수성을 보였다.

국외에서는 Hsu *et al.*(1995)은 지금까지 수문학적 예측을 위해 이용되어 온 선형적인 모형들과 신경망 이론을 적용한 모형과의 비교를 통하여 신경망의 우수성을 보여주었으며, Sajikumar *et al.*(1999)은 대상 자료의 빈약성에도 불구하고 보정·검증과정에서 좋은 결과를 나타내었다. 수질예측을 위한 인공신경망의 적용은 오창열 등(2002), 전계원(2004)이 적용한 바 있다.

그러나 일반적인 인공신경망 모형의 훈련 방법 중 최급강하법이 갖고 있는 과적합 문제와 시계반대 방향으로 전이되는 지속(Persistence)현상을 해결되지 못하였다. 이러한 문제를 해결하기 위하여 Chen *et al.*(2006)은 은닉층의 전달함수를 웨이블렛 함수로 실험을 통해 제안하였으며 그 결과 기존의 인공신경망 모형에 비하여 우수함을 보여주었다.

이러한 연구로 미루어 볼 때 지능형 모형 중 하나인 인공신경망 모형의 적용성을 위해서는 원자료에 포함된 잡음저감과 본질적인 거동 해석이 필수적 요건이라 할 수 있다.

따라서 본 연구는 비선형동역학과 웨이블렛 이론의 적용에 관한 선행 연구(오창열 등, 2007)를 통하여 최종 분해된 최적파형분해단계의 근사성분과 상세성분에 대한 잡음성분의 판별의 도구로서 인공신경망 모형의 적용 가능성을 검토하고자 하며, 더불어 인공신경망 모형의 입력자료로 활용하여 유출량 및 TOC, TOC부하량 예측모형을 개발하고자 한다.

2. 인공신경망 이론

인공신경망 모형에서의 학습이란 입력층, 은닉층, 출력층으로 구성된 다층신경망의 각 층 노드들 간의 연결강도를 최적의 상태로 적응시키는 과정을 말하며, 구조는 Fig. 1과 같다.

일반적으로 인공신경망 모형의 학습을 위한 알고리즘은 역전파학습알고리즘을 이용하며, 이에 대한 기본방법은 최급강하법이다.

최급강하법은 목적함수의 1차 도함수를 이용한 기울기에 비례하는 조정량을 산출하여 목적함수의 값이 개선될 수 있도록 매개변수의 최적화를 위해 반복적으로 탐색하는 방법이다. 본 연구에서는 최급강하법의 보다 효율적인 훈련과 더 나은 결과의 도출을 위해 모멘텀 상수와 적응식학습율을 사용하였으며 그 식은 다음과 같다.

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha[w_{ji}^{(i)}(n-1)] + \eta\delta_j^{(l)} y_i^{(l-1)}(n) \quad (1)$$

여기서, η는 학습율, α는 모멘텀 상수, $w_{ji}^{(l)}(n)$는 각 층으로 연결되는 가중치, δ는 역방향계산시 출력층 노드로 연결될 때에는 오차값과 활성화함수의 곱이며, 은닉측으로 연결될 때에는 가중치와 활성화함수의 곱이며, 본 연구에서는 모멘텀 상수와 초기 학습율은 0.1과 0.7을 적용하였다.

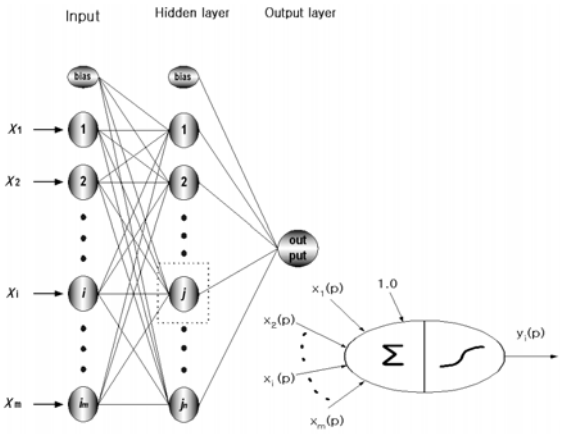


Fig. 1. General ANN Structure

3. 대상지점 및 연구방법

3.1 대상지점 및 자료

본 연구의 대상지점은 영산강 유역의 본류부를 대표하는 나주지점이며 1991년 영산강의 홍수예보 System의 일환으로 TM 수위관측소를 설치한 지점이다. 또한 환경부는 수질의 상시감시 기능과 조기경보체계의 구축을 위해 수질자동 측정망을 설치 운영 중에 있다.

Fig. 2. Geographical Location of the Naju Station

대상 자료는 2003년 1월 1일 00시부터 2004년 12월 31일 23시까지 총 17,544개의 시자료이고, 유출량은 유역내에서 발생하는 강우 등에 의하여 단위시간(sec)에 하천의 한 횡단면을 통과하는 따른 물의 양을 의미하는 것으로써 Table 1의 수위-유량관계곡선식에 시수위를 대입하여 환산한 값이다(건설교통부, 2004). 이에 따른 기술통계 분석 결과는 Table 2와 같다.

환산된 유출량 자료의 최대유출량은 2004년도 제15호 태풍 메기에 의해 7,738.01 m^3/sec가 발생하였으며 이는 계획홍수량인 5,500 m^3/sec에 비하여 2,200 m^3/sec가 초과한 양에 해당한다.

TOC는 시료 중에 용존되어 있는 유기물을 촉매산화 또는 UV산화 등의 산화과정을 통하여 CO_2가스를 발생시키고, 그 발생량을 CO_2검출기를 이용하여 전체 유기물의 양을 나타내는 수질지표이다(정명규 등, 2001). TOC 자료는 측정기기의 보정을 위한 Auto-calibration, 시약교체 및 라인세척 등으로 결측자료가 발생하게 되며, 이러한 결측자료에 대하여 본 연구에서는 일반적인 보간 방법으로 알려진 선형보간법을 이용하였다.

TOC 자료 중 최소값은 4월이며, 최대값은 6월로서 각각 1.108 mg/L, 7.891 mg/L이다. 계절적으로 홍수기 전인 6월에 최대값이 발생한 것은 영산강 수계의 산림면적이 66.9%를 차지하고 하도 내 좌·우안에 위치한 농경지에서 이루어지는 농작물 시비(施肥)와 이에 따른 강우시 비점오염물질 유입으로 다소 큰 농도를 보이는 것으로 판단된다.

TOC부하량은 유출량과 TOC 농도와의 곱과 단위환산계수 1,000으로 나누어 산출한 자료로서 단위시간에서의 부하량을 의미하며, 유출량 자료와 유사한 분포와 기술통계 값을 보이고 있다. 이는 TOC부하량 값이 유출량과 TOC와의 곱으로 산정된 자료이며, 비교적 큰 스케일을 갖는 유출량 자료의 영향으로 판단된다.

Table 1. Rating-curve Equations in the Naju Station

년도	적용범위	수위-유량관계곡선식	년도	적용범위	수위-유량관계곡선식
2003	$0.32 \le h \le 5.97$	$Q = 69.270 (h - 0.10995)^{1.96169}$	2004	$0.32 \le h \le 10.2$	$Q = 73.169 (h + 0.0653)^{2.0049}$

Table 2. Descriptive Statistics for the Data Used

		Max	Min	Mean	Standard deviation	Skewness	Kurtosis
Discharge (m^3/sec)		7,738.005	8.182	92.640	304.741	12.230	229.583
TOC concentration (mg/l)	Raw data	7.891	1.108	3.579	0.967	0.761	0.603
	Interpolated data	7.891	1.108	3.570	0.974	0.795	0.657
TOC load(kg/sec)		30.295	0.019	0.306	1.111	14.323	296.240

3.2. 예측모형의 구성 및 평가기준

인공신경망 모형은 훈련과정에서의 입력자료 범위를 초과하거나 미만인 자료에 대해서는 외삽추정에 의한 예측력이 떨어지는 특성을 갖고 있다. 따라서 훈련자료는 Fig. 3에서 나타난 바와 같이 2004년 8월 태풍 메기에 의하여 큰 유출량과 TOC부하량 값을 갖는 2004년도 자료를 이용하였으며 검증자료는 2003년도 자료를 적용하였다.

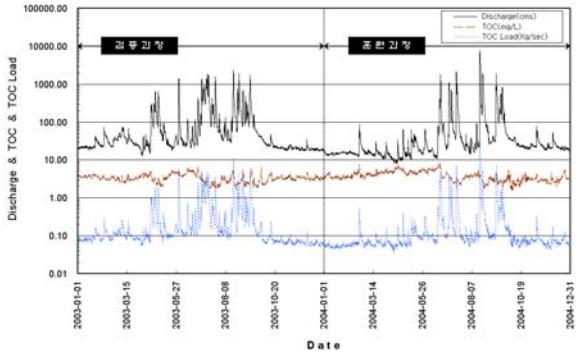


Fig. 3. Division of Input Data for Discharge, TOC, Load of TOC

더불어 인공신경망 모형의 수렴속도를 높이기 위하여 입력 자료를 Eq. (2)을 통해 정규화하였다. 기 연구된 인공신경망 모형의 정규화는 0.1과 0.9 또는 0.0과 1.0 범위로 입력 자료를 처리하였다(오창열 등, 2002; 전계원, 2004; 박성천 등, 2005; 박성천 등, 2006a; 박성천 등, 2006b). 그러나 웨이블렛 변환에 의해 최종 분해된 각 단계별 상세성분은 선행연구(오창열 등, 2007)에서 나타난 바와 같이 웨이블렛 변환의 수렴조건에 의하여 0을 중심으로 대칭형 분포를 가짐에 따라 -0.9와 0.9 사이로 정규화 하였다.

$$Y = \frac{1.8}{X_{max} - X_{min}}(X_i - X_{min}) \tag{2}$$

여기서, Y는 정규화된 자료값, X_i는 각 자료의 i번째 자료, X_{max}는 원자료의 최대값, X_{min}는 원자료의 최소값이다.

본 연구는 1시간 후 유출량 및 TOC, TOC부하량 예측 모형을 개발하기 위하여 선행연구 결과로 기 확보된 웨이블렛 변환함수 'db10'에 의하여 최종 변환된 최적파형분해단계의 근사성분과 각 상세성분을 이용하여 다음의 Eqs. (3)~(10)까지 구성하였다.

이에 대한 모형의 적용성과 예측력에 대한 평가 기준은 WMO(world meteorological organization, 1975)에서 제안한 수치적 기준과 도식적 기준을 이용하였다.

수치적 기준으로는 모형의 효율성에 대한 판단기준으로 결정계수(R^2) 값과 모형의 적합도 검증을 위하여 *RMSE*(Root Mean Square Error), 유출량 및 수질오염사고에 대한 예측을 통해 수자원의 효율적인 관리 및 재난재해에 보다 적극적으로 대응하기 위한 지표로서 *PEP*(Peak Error Percent)를 이용하였으며 이에 따른 통계적 특성치는 Table 3과 같다.

도식적 기준은 입력 자료의 구성에 따른 관측값과 모의값에 대한 시계열 곡선과 산포도를 도시하여 평가하였다.

$$Model\ I\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1) \right] \tag{3}$$

$$Model\ II\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1), D7(t,\ t-1) \right] \tag{4}$$

$$Model\ III\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1), D7(t,\ t-1), D6(t,\ t-1) \right] \tag{5}$$

$$Model\ IV\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1), D7(t,\ t-1), D6(t,\ t-1), D5(t,\ t-1) \right] \tag{6}$$

$$Model\ V\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1), D7(t,\ t-1), D6(t,\ t-1), D5(t,\ t-1), D4(t,\ t-1) \right] \tag{7}$$

$$Model\ VI\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1), D7(t,\ t-1), D6(t,\ t-1), \cdots\cdots\cdots; D3(t,\ t-1) \right] \tag{8}$$

$$Model\ VII\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1), D7(t,\ t-1), D6(t,\ t-1), \cdots\cdots\cdots; D2(t,\ t-1) \right] \tag{9}$$

$$Model\ VIII\ (Dis., TOC,\ T.L.) = f\left[A7(t,\ t-1), D7(t,\ t-1), D6(t,\ t-1), \cdots\cdots\cdots; D1(t,\ t-1) \right] \tag{10}$$

Table 3. Results of Statistics According to Each Models and Data

			Model Ⅰ	Model Ⅱ	Model Ⅲ	Model Ⅳ	Model Ⅴ	Model Ⅵ	Model Ⅶ	Model Ⅷ
Dis.	Training	R^2	0.314	0.396	0.848	0.932	0.976	0.989	0.992	0.986
		RMSE	311.579	292.228	144.553	96.761	57.298	37.577	32.696	42.794
		PEP	-87.566	-82.757	-26.228	-18.165	-4.396	1.902	-5.112	-9.702
	Verification	R^2	0.493	0.567	0.82	0.942	0.986	0.988	0.989	0.988
		RMSE	163.546	155.633	109.454	55.156	28.059	25.207	26.000	25.772
		PEP	-54.289	-18.754	-42.496	-17.648	11.043	-0.818	-5.473	0.819
TOC	Training	R^2	0.937	0.961	0.979	0.989	0.996	0.998	0.998	0.998
		RMSE	0.414	0.350	0.252	0.186	0.109	0.078	0.058	0.060
		PEP	-19.875	-12.358	-7.257	-5.255	-4.992	-3.717	-3.024	-3.960
	Verification	R^2	0.886	0.942	0.974	0.987	0.995	0.997	0.998	0.997
		RMSE	0.417	0.273	0.182	0.135	0.080	0.069	0.060	0.065
		PEP	-19.529	-16.776	-4.851	-2.259	-1.419	-1.220	1.266	1.756
TOC Load	Training	R^2	0.561	0.589	0.913	0.966	0.976	0.991	0.990	0.989
		RMSE	1.423	1.349	0.328	0.130	0.098	0.033	0.045	0.048
		PEP	-88.176	-86.856	-27.016	-18.891	-11.523	-8.787	-14.83	-17.464
	Verification	R^2	0.605	0.785	0.905	0.978	0.985	0.992	0.991	0.994
		RMSE	0.353	0.222	0.120	0.024	0.017	0.010	0.010	0.007
		PEP	-70.781	-71.243	-44.164	-7.858	-13.093	-5.524	4.839	-7.142

여기서, Dis.는 유출량, T.L.은 TOC부하량을 의미하며, A7은 각 대상 자료에 웨이블렛 변환의 적용에 따른 최종파형분해단계의 근사성분, D7부터 D1은 각 단계별 상세성분을 의미한다.

일반적으로 인공신경망 모형의 입력 자료는 대상지점에 영향을 미치는 선행강우량 및 선행유출량 자료를 자기상관분석 등을 통하여 결정한다. 그러나 본 연구에서는 서론의 연구목적에 기술하였듯이 대상자료에 포함된 잡음성분 판별하기 위하여 인공신경망을 적용하였으며, 최적의 입력자료를 구성하고자 하였다. 이는 웨이블렛 변환은 다양한 주파수 성분에 대한 유연한 분해능을 갖는 특성에 근거한다.

유출량 및 TOC부하량 예측모형에 대한 훈련 및 검증과정의 통계적 특성치를 살펴보면 Model Ⅰ부터 Model Ⅵ까지 점진적으로 개선되는 경향을 보였으나 Model Ⅶ 이후의 모형에서는 예측의 정확성이 낮아지는 경향을 보였다. 부가적으로 Model Ⅵ(Dis.) 모형은 검증과정에서 Model Ⅶ와 Model Ⅴ에 비하여 첨두유출량에 대한 예측력이 우수한 것으로 분석되었다. 또한, Model Ⅵ(T.L.) 모형의 첨두 TOC부하량은 검증과정에서 Model Ⅶ(T.L.)에 비하여 과소평가된 것으로 나타났으나, 훈련과정을 포함한 전체적인 결과에서 가장 우수한 것으로 나타났다.

TOC 예측모형은 Model Ⅰ(TOC)부터 Model Ⅶ(TOC)까지 점진적으로 개선되는 경향을 보였으나 Model Ⅶ(TOC) 이후의 모형에서는 크게 개선되지 않는 것으로 나타났다. 또한 Model Ⅷ(TOC) 모형은 Model Ⅵ(TOC) 모형에 비하여 검증과정에서 다소 과대평가하는 것으로 나타났으나 각종 수질 오염사고 대비하고자 하는 본 연구의 목적에 비추어 볼 때 Model Ⅷ(TOC) 모형이 가장 우수한 것으로 판단된다.

따라서 고주파수 성분인 상세성분 D1은 잡음성분으로 판단되며, 이러한 결과는 선행연구에서도 가장 작은 에너지 값을 갖는 것으로 나타난 바 있다.

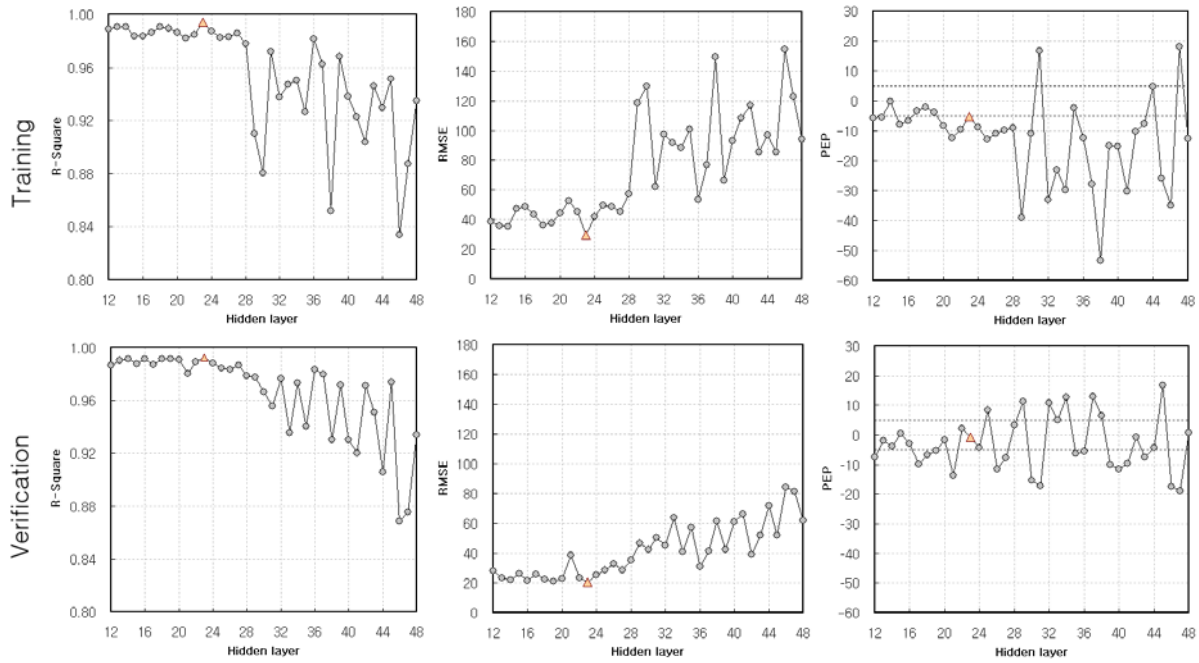
4. 모의결과 및 고찰

4.1 유출량 예측모형

일차적으로 선택된 모형인 Model Ⅵ(Dis.)에 대하여 인공신경망의 매개변수 중 하나인 은닉층 노드의 수에 따른 최적의 모형을 탐색하였으며 수치적 기준인 R2 및 *RMSE*, *PEP*값의 변화 곡선은 Fig. 4와 같다.

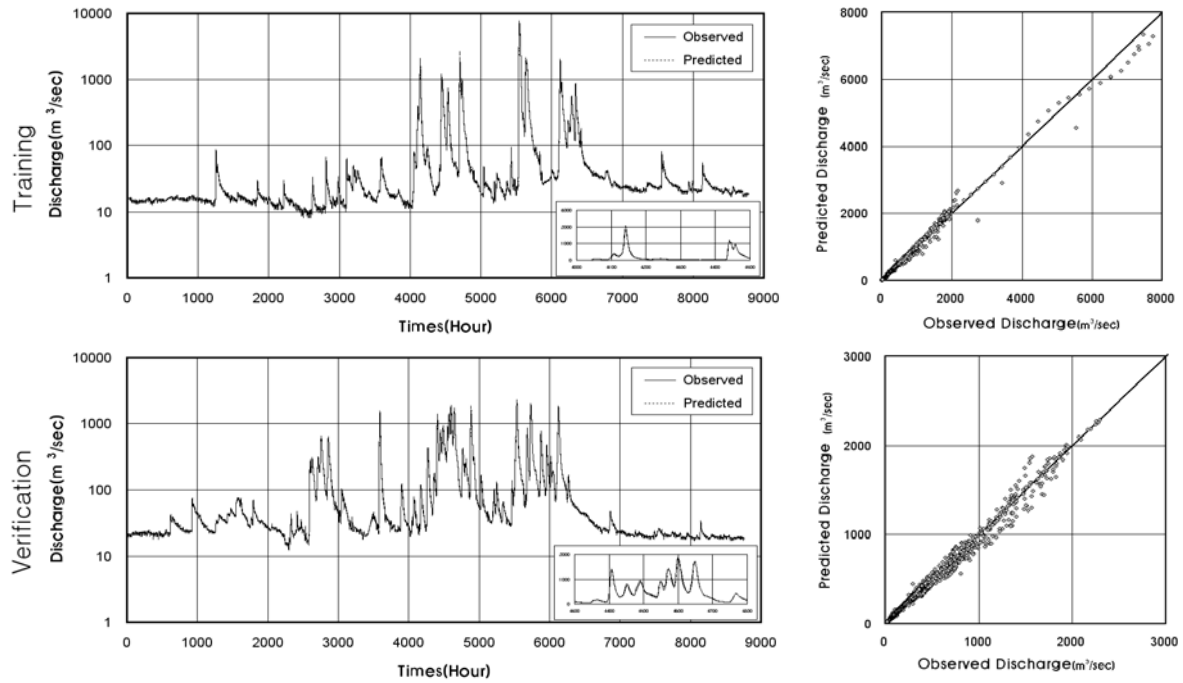
수치적 기준과 도식적 기준을 이용하여 본 연구에서 선정한 나주지점의 유출량의 특성을 가장 잘 반영하고 있는 모형은 Model Ⅵ(Dis.)_23 모형이며, 본 모형은 최종 파형 분해단계 및 D7~D3의 근사·상세성분에 대한 시차 1~2로 구성된 모형이다.

본 모형은 훈련 및 검증과정에서 가장 우수한 R^2 및 *RMSE* 통계치를 보였으며 *PEP*에 대한 통계치를 살펴보면 훈련과정에서는 다소 과소평가하고 있었으나 검증



(a) R-Square (b) RMSE (c) PEP

Fig. 4. Values of Statistics According to Hidden Layer for Training & Verification Course of Model Ⅵ(Dis.)



(a) Time series plots between raw and predicted (b) Scatter plot

Fig. 5. Plots of Verification Course for Model Ⅵ(Dis.)_23 Model

과정에서는 Model Ⅵ(Dis.)_15 모형 다음으로 우수하게 나타났다. 그러나 Model Ⅵ(Dis.)_15 모형의 R^2 및 *RMSE*에 대한 통계치를 Model Ⅵ(Dis.)_23 모형과 비교하여 볼 때 본 연구에서 선택한 모형이 확연하게 우수한 예측력을 보이고 있다.

Fig. 5(a)는 Model Ⅵ(Dis.)_23 모형에 대한 관측값과 훈련 및 검증과정에서 최적화된 연결강도에 의해 산정된 예측값을 도시한 그림이며, 산포도는 Fig. 5(b)와 같다. Figs. 5(a) and (b)에 나타낸 바와 같이 전체적으로 관측값과 예측값이 일치하고 있음을 확인할 수 있으며,

박성천 등(2006 b)이 산포도로 도시할 경우 관찰되는 지속현상은 발견할 수 없었다.

4.2 TOC 예측 모형

Model Ⅶ(TOC) 모형에 대하여 은닉층 노드의 수에 따른 최적의 모형을 탐색하였으며, 그 결과는 Fig. 6과 같다. 은닉층 노드의 수에 따른 Model Ⅶ(TOC) 모형의 R^2는 1에 가까운 0.996 이상, *RMSE* 값은 0.100 이하로서 모든 모형이 우수한 것으로 분석되었다. 반면에 *PEP* 값은 훈련과정에서는 Model Ⅶ(TOC)_14・17・19・24 모형이 좋은 예측력을 보였으나 검증과정에서는 과소평가하고 있음을 Fig. 6으로부터 알 수 있다. 이와 같은 결과를 종합하여 볼 때 본 연구에서는 훈련 및 검증과정에서의 R^2은 각각 0.998, 0.998이며 *RMSE*는 0.065, 0.056, *PEP*값은 각각 -5.130, 0.800로 나타난 Model Ⅶ(TOC)_29 모형을 선정하였다.

Model Ⅶ(TOC)_29 모형의 도식적으로 평가한 그림은 Fig. 7과 같다. Figs. 7(a) and (b)에 나타난바와 같이 전체적으로 관측값과 예측값이 일치하고 있음을 확인할 수 있으며, 본 모형에서도 유출량 예측모형과 같이 지속현상은 발견할 수 없었다. 또한 Fig. 7(b)의 산포도를 살펴보면 TOC 수질농도의 전 구간에 대하여 우수한 예측력을 보임을 직관적으로 확인할 수 있다.

더불어 각종 수질오염 사고에 대하여 보다 적극적인 대책수립을 위해서는 첨두농도에 대한 예측의 정확성이 담보되어야 하는데 본 연구에서 개발한 Model Ⅶ(TOC)_29 모형은 TOC 첨두농도에 대해서도 정확한 예측값을 보이고 있음을 Fig. 7을 통해 확인할 수 있다. 이상의 결과로 본 연구에서 개발한 Model Ⅶ(TOC)_29 모형을 수질의 상시 감시기능과 조기경보체계의 구축 업무에 적용할 경우 보다 효율적인 운용이 가능할 것으로 기대된다.

4.3 TOC부하량 예측 모형

Model Ⅵ(T.L.) 모형에 대하여 은닉층 노드의 수에 따른 통계적 특성치는 Fig. 8과 같다. TOC부하량 예측모형에 대한 통계치는 Fig. 8에서 제시된 바와 같이 본 연구의 대상모형 중 가장 큰 편차를 보이고 있다. 이는 유출량 자료에 대한 최대값과 최소값의 비율은 20배이나 TOC부하량 자료의 최대값과 최소값의 비율이 1,500배 해당하는 큰 범위를 갖는 것에 기인한 것으로 판단된다.

본 연구에서는 나주지점의 TOC부하량 자료의 특성을 가장 잘 반영하는 모형으로 Model Ⅵ(T.L.)_17 모형을 선택하였다. 수치적 기준에서는 Model Ⅵ(T.L.)_13 모형이 Model Ⅵ(T.L.)_17 모형에 비하여 우수한 결과값을 보였으나 검증과정에서는 시계반대 방향으로 전이되는 지속현상이 두드러지게 나타났다. 이는 본 연구과정에서 밝힌 바와 같이 타 연구논문과의 차별성을 갖고자 하는 본 연구목적에 부합하지 않는 것에 근거한다.

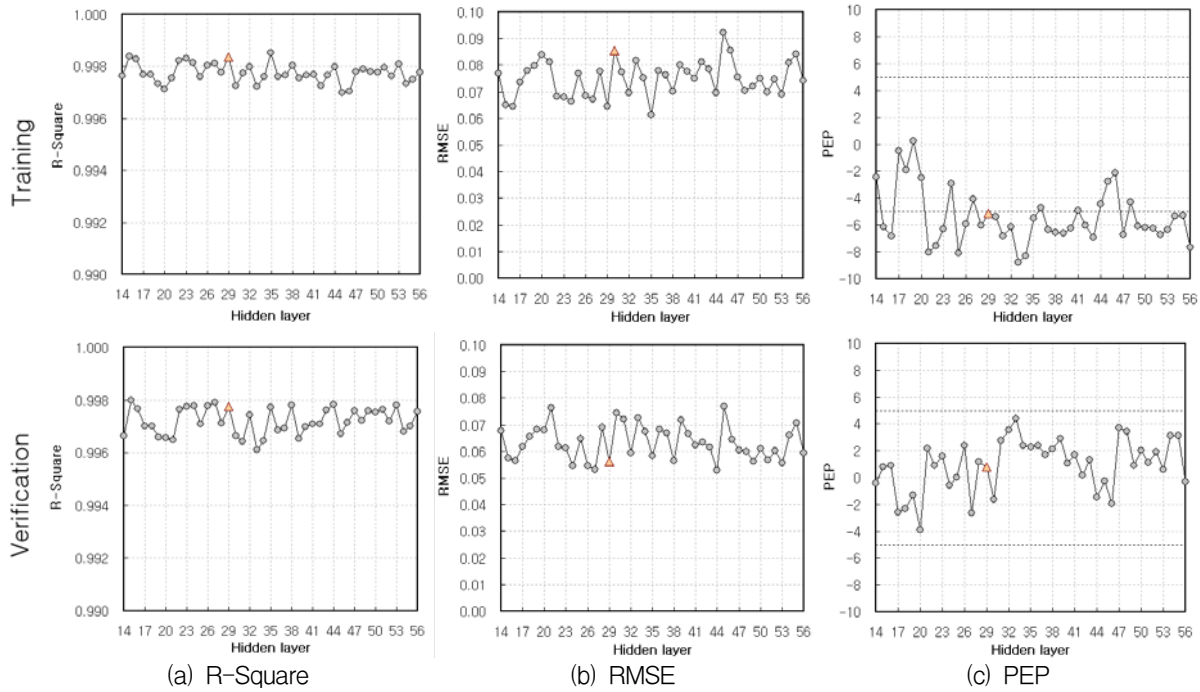
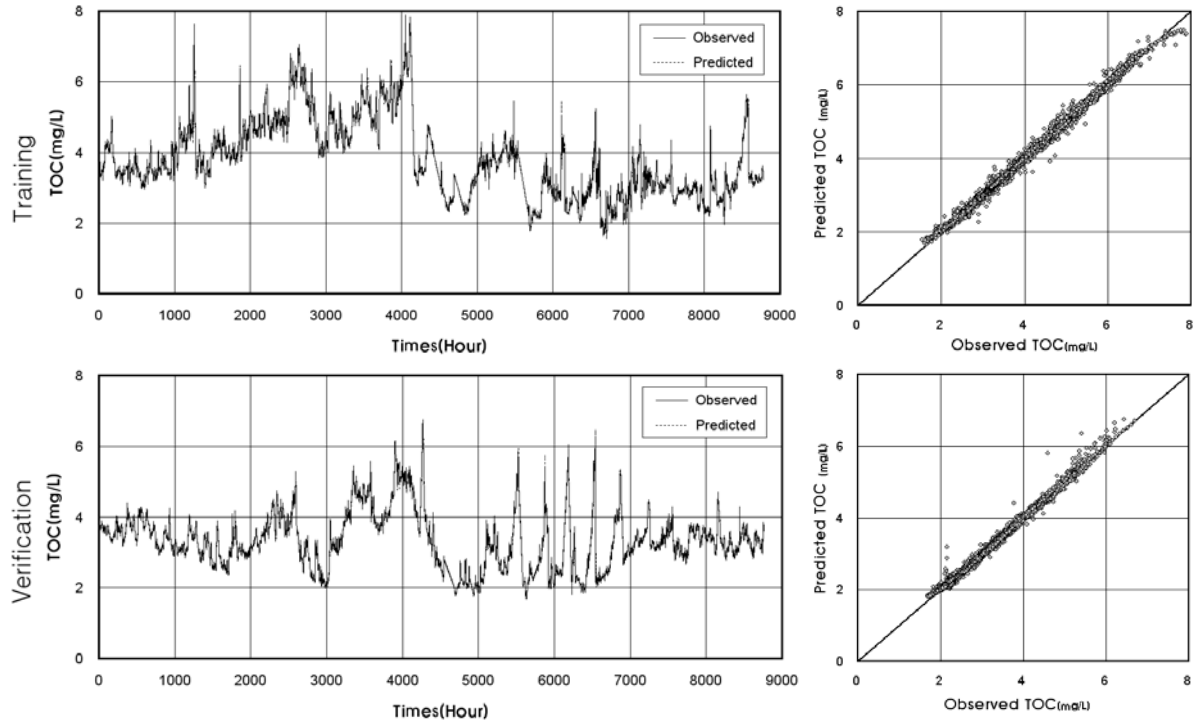
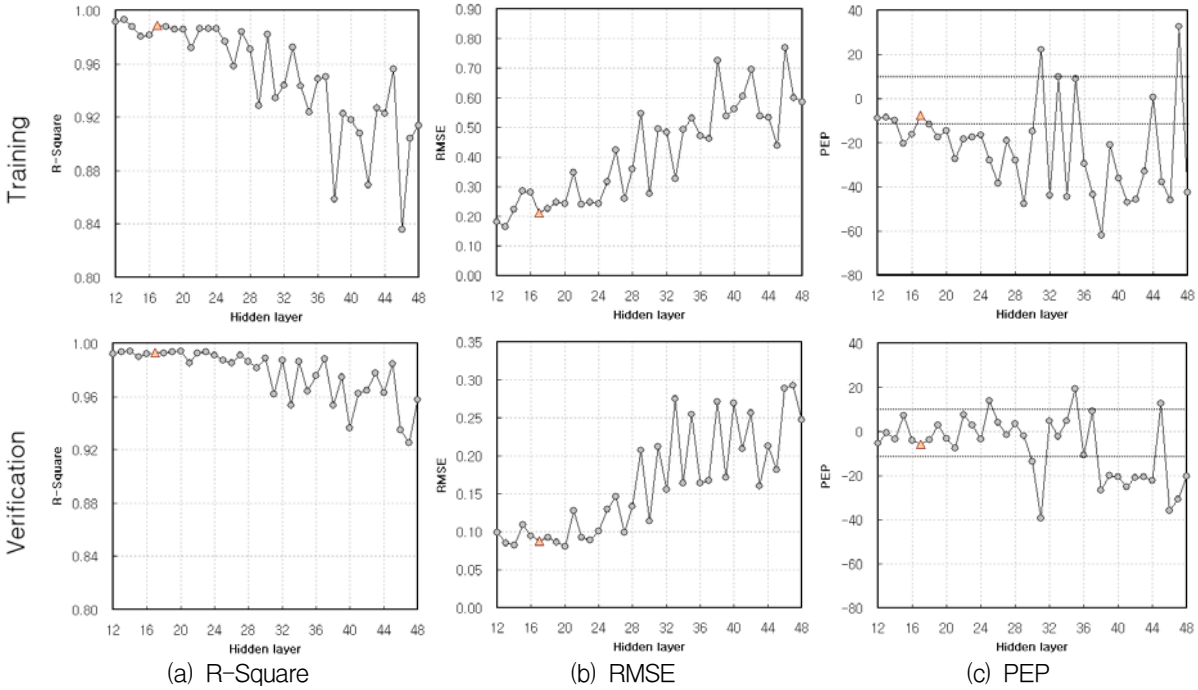


Fig. 6. Values of Statistics According to Hidden Layer for Training & Verification Course of Model Ⅶ(TOC)



(a) Time series plots between raw and predicted (b) Scatter plot

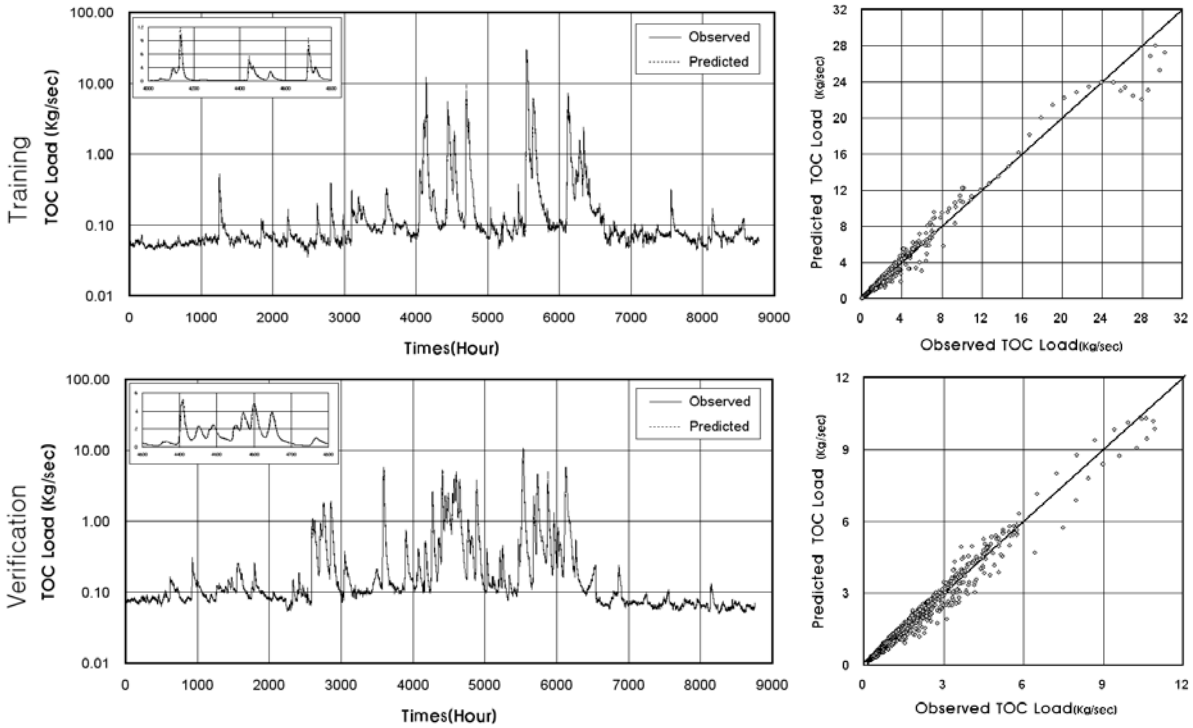
Fig. 7. Plots of Verification Course for Model Ⅶ(TOC)_29 Model



(a) R-Square (b) RMSE (c) PEP

Fig. 8. Values of Statistics According to Hidden Layer for Training & Verification Course of Model Ⅵ(T.L.)

Model Ⅵ(T.L.)_17 모형의 입력층 노드의 수는 최종 파형분해단계의 시차 1~2에 해당하는 근사성분과 D 7~D3의 상세성분으로 구성되었으며 은닉층 노드의 수는 17개로 구성된 모형이다. 훈련 및 검증과정에서의 R^2 값은 각각 0.989와 0.993이며, *RMSE* 값은 각각 0.212, 0.088이고, *PEP* 값은 각각 -7.608, -5.839의 값을 갖는 것으로 분석되었다. 전술한 바와 같이 큰 범위를 갖는 자료임에도 불구하고 R^2값과 *RMSE* 값에서는 좋



(a) Time series plots between raw and predicted (b) Scatter plot

Fig. 9. Plots of Verification Course for Model Ⅵ(T.L.)_17 Model

은 결과를 보이고 있다.

본 모형에 대한 도식적 평가는 Fig. 9(a)에 나타난 바와 같이 전체적으로 관측값과 예측값이 일치하고 있음을 확인할 수 있으며, 이를 통계적 특성치인 R^2값과 *RMSE* 값이 뒷받침하여 주고 있다. 훈련과정(Fig. 9(b))의 산포도를 살펴보면 16 Kg/sec이하의 부하량에 대한 예측값은 우수한 예측력을 보이고 있으나 16~25 Kg/sec 사이의 예측값은 다소 과대평가하고 있고, 첨두부하량까지는 과소평가하는 양상을 보였다. 또한 예측모형에 대한 검증결과 상승부에서 과소평가하는 지속현상을 다소 확인할 수 있었다.

이러한 경향은 유출량 예측모형인 Model Ⅵ_23(Dis.)과 유사한 것으로 보이고 있는데 이는 TOC부하량 자료가 유출량 자료에 많은 영향을 받는 것에 기인한 것으로 판단된다.

5. 결 론

본 연구는 유출량 및 TOC, TOC부하량 예측모형을 개발하기 위하여 잡음저감과 시간과 주파수 영역에서의 유연한 분해능을 갖는 웨이블렛 변환 적용하였으며 이에 따른 결과인 근사성분과 상세성분을 인공신경망 모형의 입력자료로 활용하였다. 더불어 최적의 예측모형 개발에 있어서 일반적으로 인공신경망에서 관찰되는 지속현상을 해결하고자 하였으며, 이러한 연구과정을 통하여 얻은 결과는 다음과 같다.

1) 본 연구에서 최종 선택한 유출량 및 TOC, TOC부하량 모형은 D1 성분을 제외한 모형으로서 D1 성분이 포함된 예측모형과 차별성을 갖는 것으로 나타났다. 이는 선행연구 과정에서도 확인하였듯이 D1 성분은 고주파수에 해당하며 가장 작은 에너지 값을 갖는 잡음성분임에 기인한 것으로 판단된다.
2) 최적의 유출량 예측모형은 Model Ⅵ_23(Dis.) 모형으로서 훈련 및 검증과정에서의 R^2에 대한 통계적 결과치는 각각 0.994, 0.993, *RMSE*는 각각 29.618, 20.454, *PEP*는 각각 -5.215, -0.647로 나타났으며 일반적인 신경망 모형에 나타나는 지속현상을 발견할 수 없었다.
3) 최적의 TOC 예측모형은 Model Ⅶ_29(TOC) 모형으로 훈련 및 검증과정에서의 R^2에 대한 통계적 결과치는 각각 0.998, 0.998, *RMSE*는 각각 0.065, 0.056, *PEP*는 각각 -5.130, 0.800로 나타났으며 본 모형 역시 일반적인 신경망 모형에 나타나는 지속현상을 발견할 수 없었다.
4) 최적의 TOC부하량 예측모형은 Model Ⅵ_17(T.L.) 모형으로 R^2에 대한 통계적 결과치는

각각 0.989, 0.993, *RMSE*는 각각 0.212, 0.088, *PEP*는 각각 -7.608, -5.839로 나타났으나 본 모형에서는 지속현상을 다소 확인할 수 있었다. 이는 대상 자료의 최대값과 최소값의 비율이 1,500배 해당하는 큰 범위를 갖는 것에 기인한 것으로 판단된다.

이상의 연구 결과는 일반적인 인공신경망 모형의 산포도에서 시계반대 방향으로 전이되는 지속현상을 극복한 것으로써 기존 인공신경망 모형에 비하여 예측의 정확도를 개선하였다는 의미를 갖는다.

본 연구의 TOC부하량 예측모형 개발로 유출량, TOC 결측자료에 대한 보간과 간접추정이 가능할 것으로 기대되며, 향후 홍수예보의 도구로서 활용하기 위하여 3시간 이상의 예측모형 개발을 위한 심화연구가 필요할 것으로 판단된다.

참 고 문 헌

건설교통부 영산강홍수통제소 (2004). **영산강 수계 유량측정보고서**. (주)한국종합기술개발공사.

김주환 (1993). **신경회로망을 이용한 하천 유출량의 수문학적 예측에 관한 연구**. 박사학위논문, 인하대학교.

박성천, 오창열, 김동렬, 진영훈 (2006a). "인공신경망 이론을 이용한 홍수유출 예측시스템 개발 -GUI-FFS 개발 및 적용-." **대한토목학회논문집**, 대한토목학회, 제26권, 제2호, pp. 145-152.

박성천, 진영훈, 김용구 (2006b). "강우-유출 예측모형 개발을 위한 자기조직화 이론의 적용." **대한토목학회논문집**, 대한토목학회, 제26권, 제4호, pp. 389-298

오창열, 박성천, 이한민, 표영평 (2002). "신경망을 이용한 영산강의 수질예측", **대한토목학회논문집**, 대한토목학회, 제22권, 제3B호, pp. 371-382.

오창열, 진영훈, 박성천 (2007). "영산강 유역의 유출량 및 수질자료에 대한 비선형 동역학과 웨이블렛 이론의 적용." **한국물환경학회지**, 한국물환경학회, 제23권, 제4호. pp. 561-560

윤강훈, 서봉철, 신현석 (2004a). "신경망을 이용한 낙동강 유역 홍수기 댐유입량 예측." **한국수자원학회논문집**, 한국수자원학회, 제37권, 제1호, pp. 67-75.

윤강훈, 서봉철, 신현석 (2004b). "신경망을 이용한 낙동강 유역 하도유출 예측 및 홍수예경보 이용." **한국수자원학회논문집**, 한국수자원학회, 제37권, 제2호, pp. 145-154.

윤강훈, 서봉철 (2004c). "댐방류량을 이용한 한강 유역 신경망 예측모형에 관한 연구." **대한토목학회논문집**, 대한토목학회, 제24권, 제1B호, pp. 47-53.

이관수, 박성천, 이한민, 진영훈 (2000). "인공신경망 이론의 B.P. 알고리즘을 적용한 영산강의 유출량 예측에 관한 연구." **대한토목학회논문집**, 대한토목학회, 제20권, 제5B호, pp. 679 - 688.

이정규, 박성제 (2003). "다층신경망을 이용한 첨두유량 예측에 관한 연구." **대한토목학회논문집**, 대한토목학회, 제23권, 제2B호, pp. 105-114.

전계원 (2004). "홍수유출과 수질예측을 위한 Web기반 시스템의 활용." 박사학위논문, 충북대학교.

정명규 외 7인 (2001). **환경분석화학**. 동화기술, pp. 406-407.

Chen, Y., Yang, B., Dong J. (2006). "Time-series prediction using a local linear wavelet neural network." *Neurocomputing*, Vol. 69, No. 4-6, pp. 449-465.

Hsu, K.L. & Gupta, H.V. & Sorooshian, S. (1995). "Artificial neural network modeling of the rainfall-runoff process." *Water Resources Research*, Vol. 31, No. 10, pp. 2517-2530.

Sajikumar, N. & Thandaveswara, B. S. (1999). "A non-linear rainfall-runoff model using an artificial neural network." *Journal of Hydrology*, Vol. 216, pp. 32-55.

World Meteorological Organistaion (1975). *Inter-comparison of conceptual models used in operational hydrological forecasting.* World Meteorological Organisation, Techinical report No 429, Geneva, Switzerland

(논문번호:08-75/접수:2008.06.25/심사완료:2008.09.03)