

신경망 모형을 이용한 달천의 수질예측 시스템 구축

Construction of System for Water Quality Forecasting at Dalchun Using Neural Network Model

이원호¹ · 전계원^{2,*} · 김진극³ · 연인성⁴

Lee, Won-ho¹ · Jun, Kye-won^{2,*} · Kim, Jin-geuk³ · Yeon, In-sung⁴

1 충주대학교 토목공학부 도시공학전공

2 강원대학교 방재기술전문대학원

3 하이드로시스넷

4 충북대학교 건설기술연구소

(2007년 2월 2일 논문 접수; 2007년 6월 14일 최종 수정논문 채택)

Abstract

Forecasting of water quality variation is not an easy process due to the complicated nature of various water quality factors and their interrelationships. The objective of this study is to test the applicability of neural network models to the forecasting of the water quality at Dalchun station in Han River. Input data is consist of monthly data of concentration of DO, BOD, COD, SS and river flow. And this study selected optimal neural network model through changing the number of hidden layer based on input layer(n) from n to 6n. After neural network theory is applied, the models go through training, calibration and verification. The result shows that the proposed model forecast water quality of high efficiency and developed web-based water quality forecasting system after extend model

Key words: neural network, water quality, web-based, forecasting

주제어: 신경망, 수질, 웹기반, 예측

1. 서 론

최근 이상기후와 도시화, 산업화 및 농업의 근대화 등으로 수질오염 물질의 배출이 광역화 되고 있으며, 이로 인해 수질개선을 위한 연구와 기술 개발이 활성화 되고 있다.

유량과 수질 인자들간의 상호관계, 다양한 인자들이 수질변화에 미치는 영향 등을 분석하고자 하는 노력이 이어져 하천 수질의 예측 및 모의를 위해 많은 모형들과 새로운 방법론이 제시되어 왔다.

일반적인 수질예측 및 분석에는 많은 매개변수를 요하게 되며, 결과에 영향을 미치는 불확실성을 제거하기 위한 노력의 일환으로 최근 인공지능 알고리즘

*Corresponding author Tel: +82-33-570-6816, FAX: 82-33-570-6819, E-mail: kwjun@kangwon.ac.kr (JUN, K.W.)

을 수질예측에 적용하려는 시도가 있어 왔다. 인공지능 알고리즘은 인간의 사고과정을 모방하여 수치적 모의를 제시한 것으로서 자료들의 단순관계 해석을 넘어 비선형적 상관성을 해석하기에 적합하며, 이를 모형화한 대표적인 것으로 신경망 모형을 들 수 있다 (연인성, 2005).

인공신경망은 적용이 간편할 뿐 아니라 공정의 비선형적인 특성을 잘 모사할 수 있고, 예측 능력이 매우 뛰어나서 최근 많이 사용되고 있다. Maier와 Dandy(1996년)는 오스트레일리아 남부에 있는 Murray강의 수질을 예측하기 위해 신경망 이론을 적용한 바 있다. Nohh(1996)는 도시하수의 수질을 모형화 하기위해 사용한 SWMM모형의 최적 매개변수 산정에 신경망 이론을 적용하였고 Ken-ichi(1997)는 호수의 부영양화 예측에 BP알고리즘을 적용하고 실측치와 비교하였다.

최근 국내에서도 인공지능을 이용한 연구들이 활발히 진행되고 있다. 신 등(1998)이 도시유역 비점오염원에 대한 오염물 배출 모형화 연구를 위해 신경망 모형을 적용하였고, 이 등(2001)은 호수의 수질예측을 위해 신경망의 BP알고리즘을 적용한 바 있다. 오 등(2002)은 영산강 유역 나주지점 수질예측에, 그리고 전 등(2004)은 금강유역의 수질예측에 신경망 모형을 적용한 바 있다.

수계(水系)내에서의 수질 변동은 계절별 특성 및 수질 오염원 등에 연동하는 성질이 있어 중·장기적인 수질 예측을 위하여 신경망 이론에 의한 자기학습이 가능한 수질 모형이 요구된다. 따라서 본 연구에서는 수질예측을 위한 시스템의 개발을 위해 신경망 기법을 활용하여 한강유역의 지류인 달천지점의 수질을 예측할 수 있는 지능형 모형을 구축하고 그 적용성을 검증한 후 기초적인 웹기반 수질예측 모형을 개발하여 하천의 중·장기 수질보존 대책 수립에 기초 자료를 제공하고자 한다.

2. 신경망 모형

신경망은 은닉층의 유무에 따라 단층신경망과 같은 선형연상기억(linear associative memory)모형과 다층신경망과 같은 비선형 연상기억(nonlinear associative memory)모형으로 구분한다. 선형연상기억

모형은 일반적으로 학습패턴간의 선형적 특성에 따라 네트워크의 수렴여부가 좌우되며, 비선형 연상기억모형은 학습패턴간의 비선형 분리능력이 있고 임의의 복잡한 영역의 분류가 가능하고 정보의 기억능력도 매우 크다.

신경망은 하나 또는 다수의 층들로 이루어지는데 하나의 층으로 이루어진 것을 단층 신경망이라고 하고 다수 개의 층들로 이루어진 신경망을 다층신경망이라고 한다. 다층 구조의 신경망에서 각 층사이의 연결강도는 모형에 따라 미리 계산에 의해 주어지는 경우와 학습에 의해 결정되는 경우가 있으며, 본 연구에서는 각각의 입력에 대해 기대하는 출력과 동일하거나 유사한 출력을 얻기 위해 반복학습을 통하여 신경망의 연결강도의 값을 구하는 학습(learning)과정을 사용하였다. 또한 이러한 학습과정을 위해 역전파 알고리즘의 형태에 속도개선과 불안정한 학습결과를 개선하기 위해 모멘트법(moment)과 적응학습률(adaptive learning rate) 방법을 사용하였다. 모멘트법은 수렴 속도를 증진시켜 학습의 고숙화에 유효한 것으로 알려져 있으며 간단하면서도 효율적이기 때문에 일반적으로 많이 사용되고 있다. 기본방정식은 식 (1)과 같이 나타낼 수 있다.

$$\begin{aligned} \Delta W_{jk}(t_n + 1) &= \eta \delta_{pk} b_{pj} + \beta_1 \Delta W_{jk}(t_n) \\ \Delta W_{ji}(t_n + 1) &= \eta \delta_{pj} b_{pi} + \beta_2 \Delta W_{ji}(t_n) \end{aligned} \quad (1)$$

여기서, t_n 는 반복횟수, δ_{pk} , δ_{pj} 는 처리소자의 오차, b_{pi} , b_{pj} 는 은닉층의 출력값, η 는 학습율이다.

적응학습율은 모멘트법과 병행하여 적응학습율 η 를 조절하며, 학습 속도를 향상시키는 방법으로 기본식은 식 (2)와 같이 나타낼 수 있다.

$$\eta(t+1) = \begin{cases} r_1 \eta(t), & E(t+1) < E(t) \\ r_2 \eta(t), & E(t+1) > r_3 E(t) \\ \eta(t), & otherwise \end{cases} \quad (2)$$

여기서, r_1 , r_2 , r_3 는 적응학습율에 따른 매개변수, $E(t+1)$, $E(t)$ 는 현 단계 및 전 단계의 오차를 나타낸다.

3. 수질예측 모형의 적용

3.1. 적용자료

수질자료는 신경망 모형의 학습 및 예측을 위해 필요하며 본 연구에서는 분석에 사용된 수질자료는 월 대표수질을 사용하였다. 수온, DO, BOD, SS 자료는 1994년 1월부터 2003년 12월까지 10년간의 월 대표수질 자료를 사용하였고 선택된 수질자료를 분석하여 각 지점의 수질변화 특성을 파악하였다. DO의 월 평균에 대한 1월부터 12월까지의 평균값은 8.02~13.05mg/L로 분석되었으며, BOD는 0.92~1.77 mg/L, SS는 1.82~9.56mg/L로 분석되었다.

본 연구에서 사용된 수질자료 중 1994년부터 2001년까지의 자료는 신경망 모형의 연결강도 산정 및 검증자료로 사용되었으며 2002년 수질자료는 최적모형의 결정을 위한 검증자료로 사용되었다. 그리고 2003년 수질자료는 예측시스템 구축시 Test 자료로 활용하였다.

3.2. 수질예측 모형의 구성

남한강유역 달천지점의 수질예측을 위해서 각 지점별 수온, DO, BOD, COD, SS 등의 수질인자와 월 유량 자료를 이용하여 하천수질예측 모형을 구성하였다.

하천수질예측의 모형화를 위해 시간의 흐름에 따른 비선형 수질예측 인자들의 관련성은 BOD 예측모형을 통하여 살펴보면 식 (3)과 같이 정의할 수 있다.

$$BOD(t+1) = f[BOD(t-k_1), Q(t-k_2)]$$

$$\text{for } k_1 = 0, \dots, n, k_2 = 0, \dots, m \quad (3)$$

여기서, $BOD(t+1)$ 은 1시간 후의 예측 BOD이고, $Q(t-k_1)$ 는 k_1 시간 전의 유량, $Q(t-k_2)$ 는 k_2 시간 전의 유량, k_1, k_2 는 각각 n, m 개의 시계열이다.

구성된 예측모형은 DO, SS모형 구성시에도 활용되며 입력자료의 형태만 수질예측인자에 따라 변화시켜 예측 모형을 구성하였다.

3.3. 학습을 통한 모형의 구축

모형의 구성단계에서 구축된 신경망 모형을 대상 지점에 적용하여 학습을 수행하고 각 지점에 적합한 연결강도를 찾아 예측을 위한 모형의 검정을 하였다. 달천지점의 1994년부터 2001년까지의 수질자료를 학습시켜 각 모형에 대한 최적의 연결강도를 구한 후 2002년 1월부터 12월까지의 수질자료를 이용하여 모형의 예측능력을 검증하였다.

신경망 모형을 훈련하기 위해 입력자료의 수집 및 분류가 이루어진 후 목표오차에 이르는 훈련을 위해 충분한 은닉 뉴런의 개수를 갖는 네트워크의 구조를 선택하고 연결강도 가중치와 임계값들을 초기화 시킨 후 훈련을 수행하였다. 학습시 신경망의 구조가 목표 오차에 도달되는 가장 작은 네트워크가 될때 학습이 종료되며 선택된 목표오차를 만족하는 네트워크가 발견되면 검증과정을 수행하게 된다.

출력자료에 대한 입력자료의 지체시간에 따른 항목별 상관관계는 수질 항목별로 차이는 있으나 대부분 달천지점의 1개월 전의 자료가 현재의 달천지점의

Table 1. 수질 예측을 위한 신경망 모형의 구성

| 항목 | 모형 | 입력층 노드수(N) | 은닉층 노드수 | 입력자료 | 출력자료 |
|-----|------|------------|---------|---------------------|--------------|
| 수온 | MANN | 1 | 2N~6N | $TEMP_t$ | $TEMP_{t+1}$ |
| | | 2 | 2N~6N | $TEMP_t, DO_t$ | $TEMP_{t+1}$ |
| DO | MANN | 1 | 2N~6N | DO_t | DO_{t+1} |
| | | 2 | 2N~6N | $DO_t, TEMP_t$ | DO_{t+1} |
| BOD | MANN | 1 | 2N~6N | BOD_t | BOD_{t+1} |
| | | 2 | 2N~6N | BOD_t, COD_t | BOD_{t+1} |
| | | 3 | 2N~6N | BOD_t, Q_t, COD_t | BOD_{t+1} |
| SS | MANN | 1 | 2N~6N | SS_t | SS_{t+1} |
| | | 2 | 2N~6N | S_t, Q_t | SS_{t+1} |

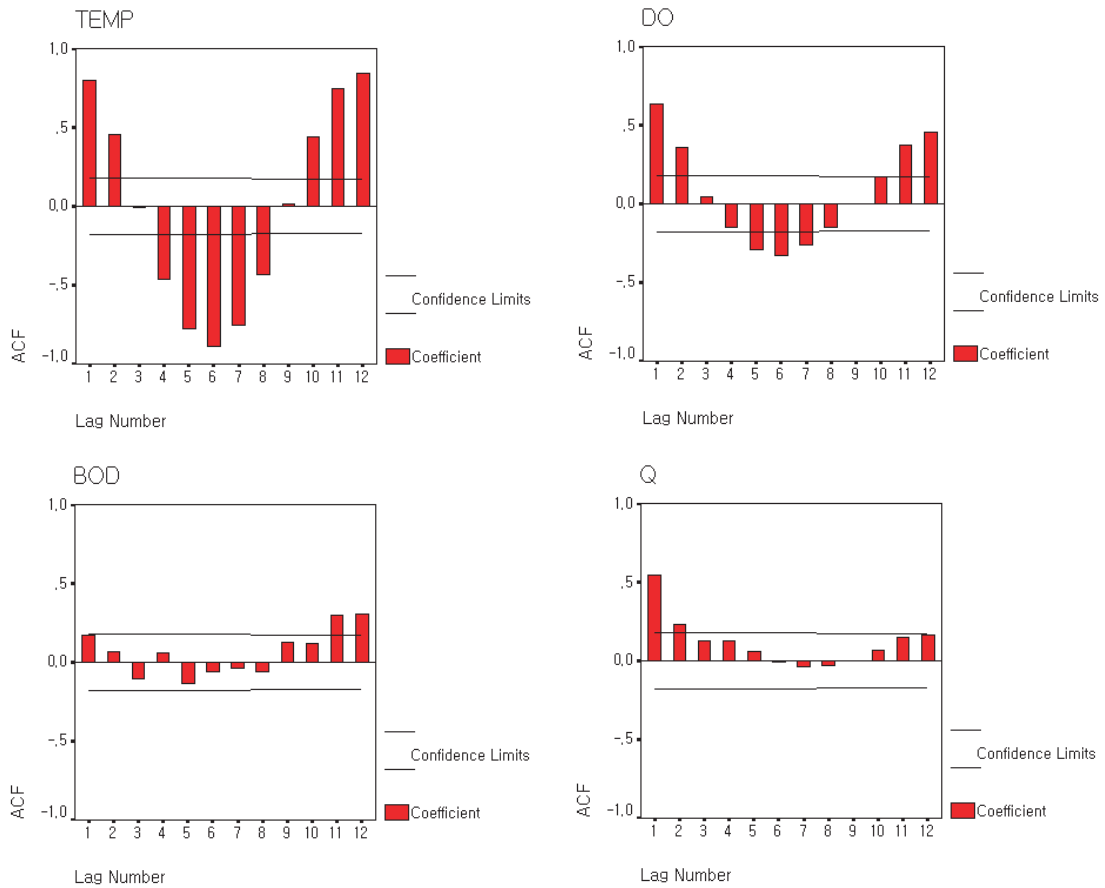
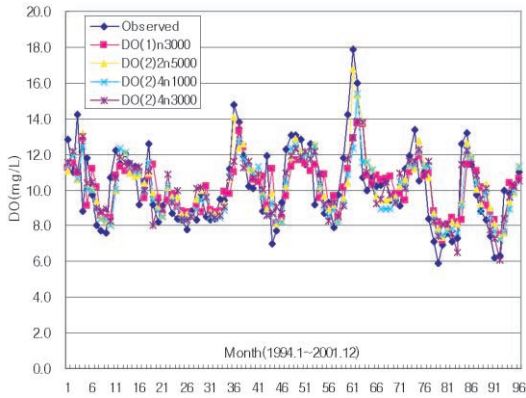


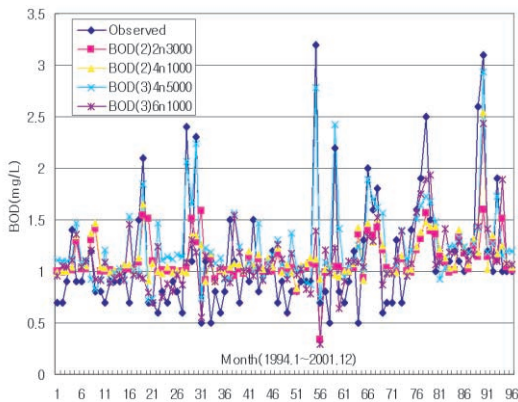
Fig. 1. 수질인자에 대한 자기상관 분석.

Table 2. 달천지점의 수질예측 모형의 선정을 위한 학습결과

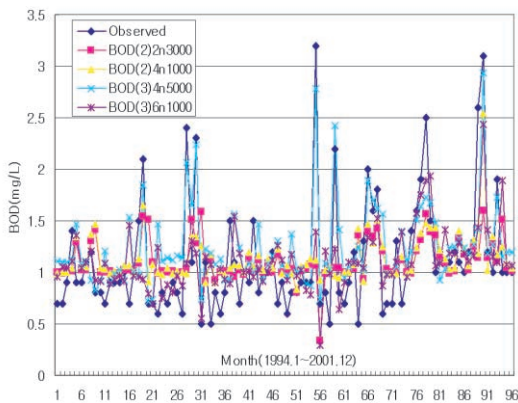
| Index | Model | Average | Standard deviation | Skewness |
|-------|--------------|---------|--------------------|----------|
| DO | Observed | 10.156 | 2.217 | 0.654 |
| | DO(1)1n3000 | 10.156 | 1.619 | 0.367 |
| | DO(1)6n1000 | 10.013 | 1.723 | 0.409 |
| | DO(2)2n5000 | 10.160 | 2.138 | 0.592 |
| | DO(2)3n3000 | 10.173 | 1.977 | 0.569 |
| BOD | Observed | 1.118 | 0.548 | 1.825 |
| | BOD(1)4n5000 | 1.107 | 0.493 | 1.512 |
| | BOD(2)2n3000 | 1.155 | 0.385 | 1.347 |
| | BOD(3)4n5000 | 1.125 | 0.512 | 1.796 |
| | BOD(3)6n5000 | 1.131 | 0.598 | 1.901 |
| SS | Observed | 5.178 | 5.754 | 5.049 |
| | SS(1)2n3000 | 4.523 | 3.276 | 7.254 |
| | SS(1)4n3000 | 4.915 | 5.174 | 6.029 |
| | SS(2)4n3000 | 5.293 | 5.375 | 5.325 |
| | SS(2)5n5000 | 5.572 | 5.412 | 6.724 |



(a) DO



(b) BOD



(c) SS

Fig. 2. 달천지점 수질예측을 위한 신경망의 학습결과.

자료와 가장 높은 상관관계를 보였으며 Lag-1에서 수온과 DO는 0.80, 0.71을 나타내며 12개월의 뚜렷한 주기성을 가지고 있음을 알 수 있다. BOD는 일정한

특성은 있으나 지체시간이 증가할수록 낮은 상관관계를 나타냈다. 수온에 대한 각 항목별 교차상관 관계를 분석한 결과 수온과 교차상관 관계가 가장 크게 나타난 항목은 DO이다. BOD의 경우 각 항목중 유량이 BOD 예측시 가장 고려해야 할 항목으로 분석되었다. 특히 유량의 경우 Lag-1의 유량이 BOD 예측시 Lag+2까지 영향을 줌으로 BOD 예측시 유량을 입력변수로 고려하였다.

달천지점의 수질예측 모형의 선정에 대한 학습결과를 Fig. 2와 Table 2에 나타내었으며 각 수질항목별로 학습정도가 우수한 모형을 수질예측을 위한 기본 모형으로 선정하였다. 특히 하천에서 수질예측에 가장 중요한 인자중에 하나인 BOD 예측의 경우 기본구성 모형을 4가지로 구성하였으며 그중 BOD 전시간 자료와 유량 전시간 자료를 사용하는 BOD(3)4n5000 모형이 예측을 위해 가장 적합한 모형으로 선정되었다. DO는 예측 결과 DO(2)2n5000모형이 최종 모형으로 선정되었으며, SS는 평균, 표준편차, 왜곡도 등을 비교해 볼 때 전달의 SS농도와 전달의 유량자료를 입력자료로 하는 SS(2)4n3000 모형이 가장 좋은 예측결과를 나타내었다.

3.4. 신경망 모형의 수질예측능력 검토

달천지점에서 수질항목별로 구성된 신경망 모형에 의해 신경망의 학습조건을 달리하여 모의 후 결정된 신경망 모형의 선정모형에 의해 2002년 1월~12월의 월 수질자료를 예측하고 수질예측 모형으로서의 예측능력을 검토하였다. Fig. 3은 학습과정에서 결정된 적정 연결강도를 갖는 신경망 모형에 의한 수질예측 결과를 나타내고 있다. DO예측의 경우 DO(2)2n5000모형, BOD예측의 경우 BOD(3)4n5000모형이 예측력이 우수한 것으로 결정되었다.

4. 수질예측 시스템의 구축

4.1. 기본시스템의 구성

예측력이 검증된 수질예측모형을 이용하여 수질예측을 위한 기본시스템을 구축하였다. 기본시스템은 클라이언트-서버 모델(Client-Server Model)로 개발되었다. 클라이언트에서 사용자가 정보를 입력하고 수행할 명령을 서버에 전달하게 된다. 서버는 전달받은

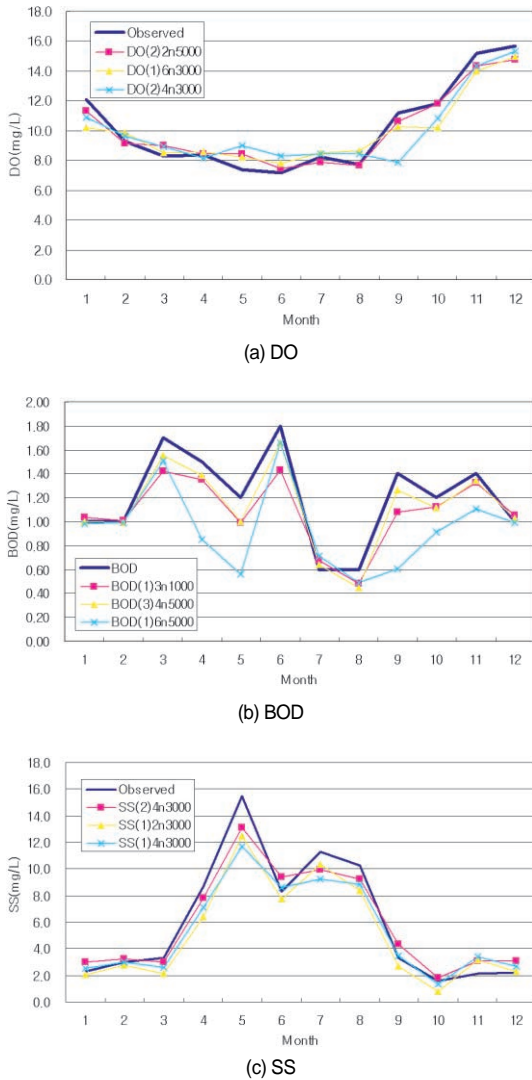


Fig. 3. 달천지점의 수질예측 결과.

명령과 정보에 따라 작업을 수행하고 결과를 클라이언트에 전송한다. 클라이언트는 웹 브라우저를 통해 최종 처리 결과를 사용자에게 표시한다. Fig. 4는 기본 시스템의 서버-클라이언트 모델을 나타낸다.

4.2. 수질예측 시스템의 구현

4.2.1. 프로그램 설계

WFS(Water quality Forecasting System)는 크게 여섯 개의 모듈(module)로 구성되며 독자적인 모듈로 수행된다(DISCHARGE, DO, BOD, SS, TN, TP). 모

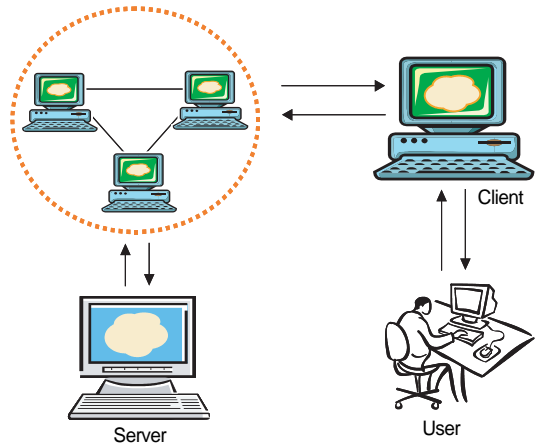


Fig. 4. 수질예측을 위한 서버-클라이언트 모델.

듈이 수행되는 시점은 WFS 웹프로그램에서 결정한다. 여섯 개의 모듈 중 DISCHARGE 모듈은 유량에 관련된 모듈이고 나머지는 수질에 관련된 모듈이다. 각각의 모듈은 학습(training), 검증(verifying) 및 예측(forecasting) 기능을 갖고 있으나 검증 기능은 예측 기능과 비슷하여 별도로 설계하지 않는다. 또한 각각의 모듈은 신경망의 모형과 입력 데이터는 다르지만 처리순서나 신경망 구성 방식이 동일하므로 하나의 공통된 모듈로 간주하고 프로그램을 설계한다. Fig. 5에는 WFS 프로그램의 최상위 설계를 나타내었다.

4.2.2. 시스템의 구현

수질예측모듈의 경우 각 지점별로 1년 단위로 예측을 수행할 수 있도록 구성하였으며 Fig. 6에 예측 모듈의 기본화면을 나타내었다. 또한 실시간 자료를 확인하고 획득하기 위한 기본모듈을 구성하고 수집된 자료들은 데이터베이스에 저장되고 사용자가 원하는 기간의 홍수량, 강우 및 수질자료의 기간을 선택하면 Web서버가 반응하여 데이터베이스에서 원하는 정보를 화면상에 표시해주는 형태로 구성하였다. Fig. 7은 사용자의 요구에 의해 수질자료를 데이터베이스에서 불러들여 Web상의 화면에 나타낸 것이다. 추후 유관기관의 실시간 데이터관리 서버와 연계된다면 개발된 시스템 상에서 실시간 자료의 효율적인 관리가 가능하다고 판단된다.

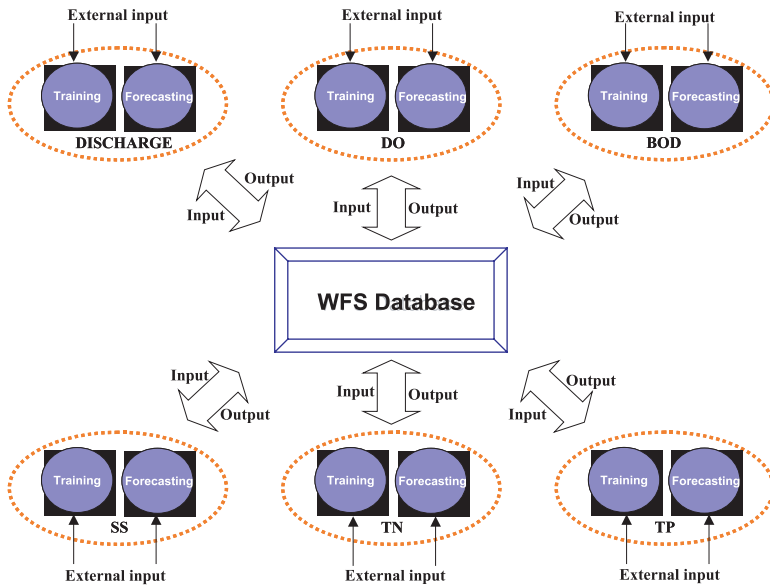


Fig. 5. WFS 프로그램의 설계.

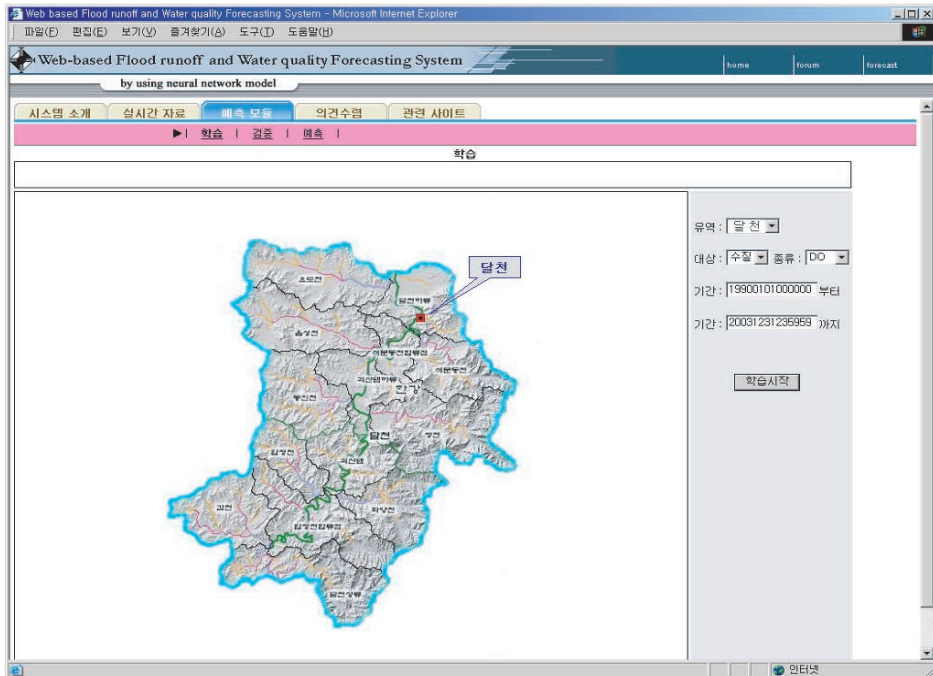


Fig. 6. 예측 모듈의 기본화면.

4. 3. 수질예측시스템의 시험

수질예측 모듈의 경우 각 수질인자에 대한 학습,

검증 및 예측과정을 수행할 수 있다. 달천지점의 수질자료를 이용하여 학습 및 검증 후 장래예측을 수행하였다. Fig. 8은 수질예측모듈의 학습수행 결과를 나



Fig. 7. 수질 자료의 검색.

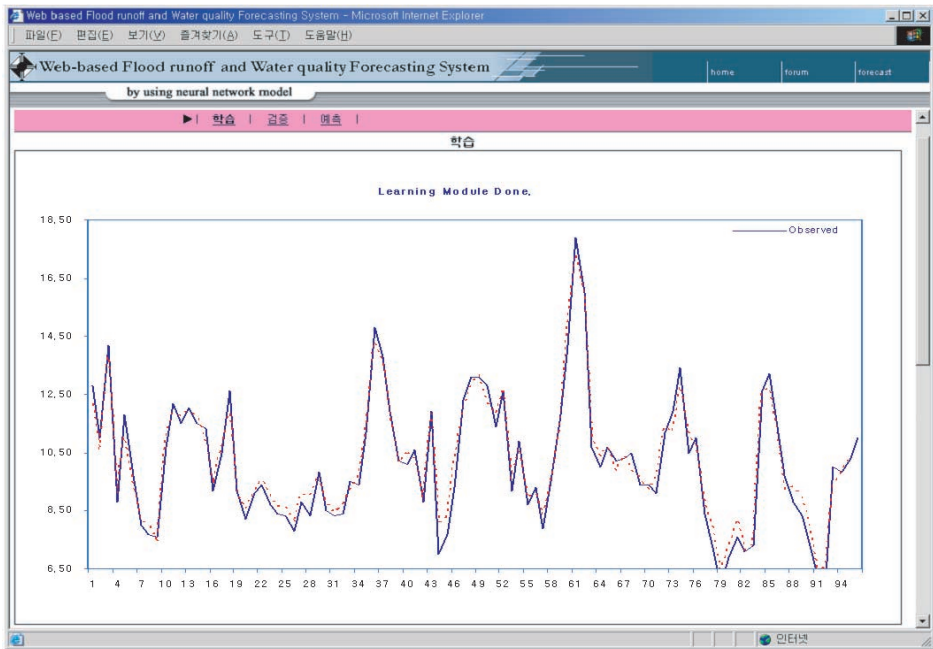


Fig. 8. 수질예측 모듈의 학습수행 결과.

타내고 있으며 신경망 모형의 학습 후 가중치 테이블에 연결강도가 저장된다. Fig. 9는 저장된 연결강도가 가중치 테이블에서 Loading 후 2002년 1월~2003년 12월까지의 달천지점의 수질예측을 수행한 검증 결과를 나타낸다. Fig. 10은 2004년 1월부터

2006년 6월까지 달천지점의 장래 수질예측을 수행한 결과를 나타내고 있다. 장래 수질예측의 경우 수질데이터가 유관기관의 Web 서버와 같은 형태로 연동된다면 계속적인 예측이 가능하다.

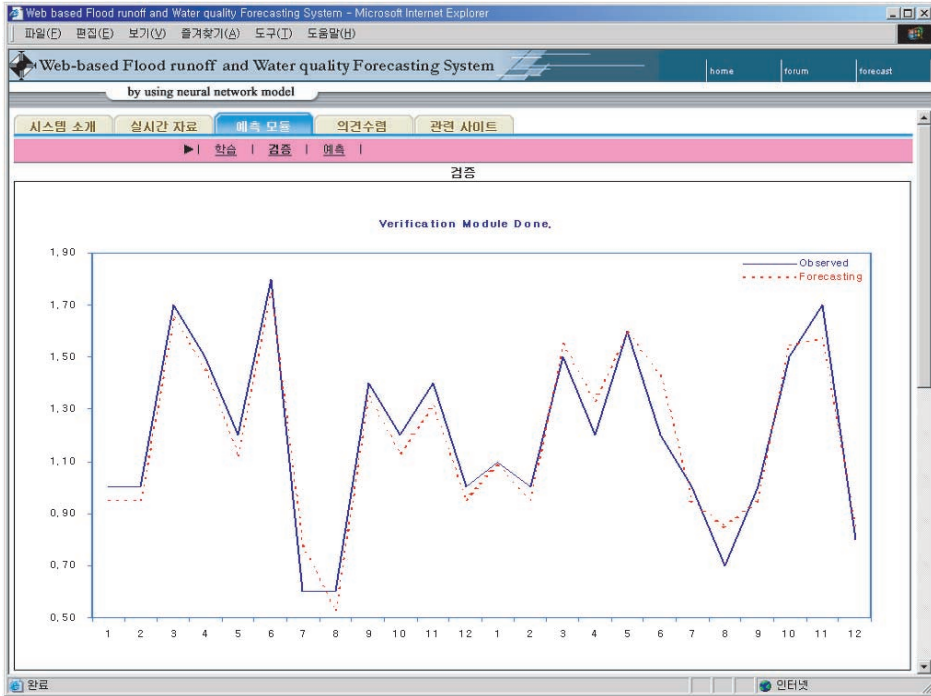


Fig. 9. 수질예측 모듈의 검증결과(BOD).

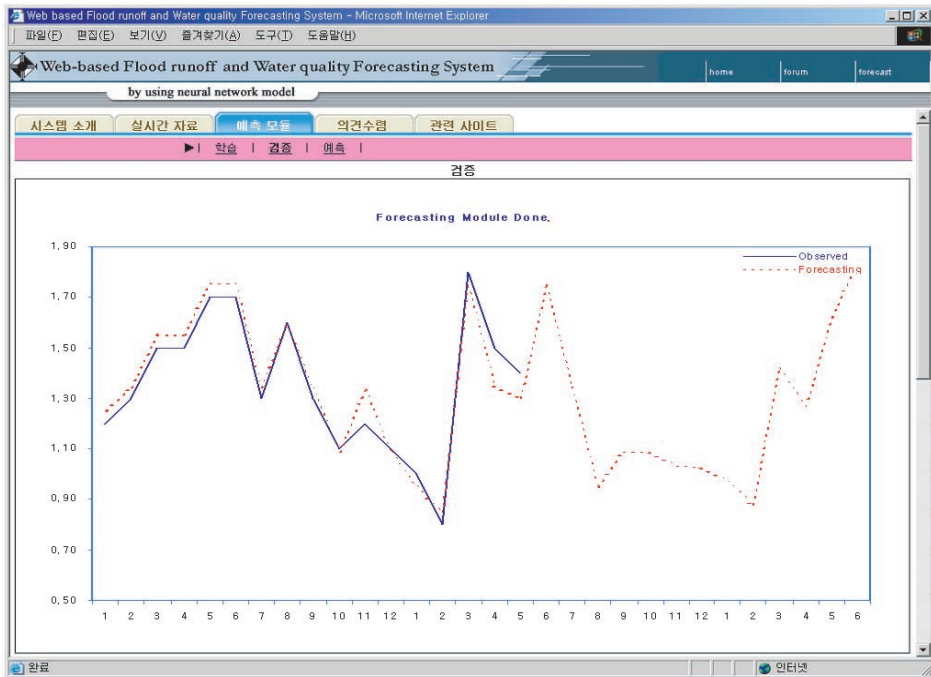


Fig. 10. 수질예측 모듈의 장래수질예측 결과(BOD).

5. 결 론

신경망 모형을 이용하여 한강수계 달천지점에 수질예측 시스템을 구축하고 그 적용성을 파악하기 위해 신경망의 기본이론, 모형의 구축 방법 및 예측시스템의 구성요소 등을 분석에 관해 수질을 Web상에서 예측할 수 있는 Web 기반 수질예측 시스템을 개발하고 Web상에서 한강수계 달천지점의 수질예측이 가능하도록 기본적인 시스템을 구성하였다. 본 연구를 통해 얻어진 결과는 다음과 같다.

(1) 한강수계 달천지점의 수질자료를 수집하여 각 항목별 수질특성을 분석하였다. 용존산소 및 수온의 경우는 계절적 특성을 반영하여 뚜렷한 주기성을 나타내었으며 BOD와 SS는 일정한 특성은 있으나 용존산소와 수온과 같은 뚜렷한 주기성은 나타나지 않았다.

(2) 수질특성 분석결과를 토대로 각 지점의 수질예측을 위해 신경망 모형을 이용하여 모형의 입력 및 출력 형태를 구성하였다. 신경망 모형의 학습에 따른 수질예측 결과 DO예측모형은 $DO(2)2n$, BOD예측모형은 $BOD(3)4n$, SS예측모형은 $SS(2)4n$ 으로 결정되었다.

(3) 수질예측을 위한 Web기반 예측시스템은 서버-클라이언트 모델로 개발되었으며 수질예측 모듈의 경우는 월 수질 예측이 가능하도록 지점별, 수질인자별 학습, 검증이 가능하며 장래예측을 수행할 수 있도록 구성하였다. 개발된 Web기반시스템에 의해 Web상에서 구현된 수질예측 모듈을 한강수계 달천지점에 적용하여 학습 및 예측의 적합성을 확인하였다.

감사의 글

본 연구는 충북지역환경기술개발센터의 2005년 연구비지원에 의해 수행되었습니다.

참고문헌

1. 신현석, 최시중, 김중훈 (1998) 신경망을 이용한 도시유역 유출 및 비점원 오염물 배출 모형화 연구. *대한토목학회논문집*, **18(5B)**, pp. 437-448.
2. 연인성 (2005) 실시간 유량-수질관리를 위한 인공지능 시스템 개발. 충북대학교 대학원박사학위논문. pp. 15-40.
3. 오창렬, 박성천, 이한민, 표영평 (2002) 신경망을 이용한 영산강의 수질예측. *대한토목학회논문집*, **22(3B)**, pp. 371-382.
4. 이경훈, 김진모, 문병석 (2001) BP 알고리즘을 이용한 호소수 수질예측. *대한상하수도학회 공동추계 학술발표회논문집*, pp. 203-206.
5. 전계원, 안상진 (2004) Web기반 홍수유출 및 수질예측 시스템의 개발 - I. 모형의 구축. *대한토목학회논문집*, **24(4B)**, pp. 321-331.
6. Ken-ichi, Y., Masaaki, H. and Akihiho, M. (1997) Novel Application of a Back-propagation Artificial Neural Network Model Formulated to Predict Algal Bloom, *Water Science and Technology*, **36**, pp. 89-97.
7. Maier, H.R. and Dandy, G.C. (1996) The Use of Artificial Neural Networks for the Prediction of Water Quality Parameters, *Water Resources Research*, **32(4)**, pp. 1013-1022.
8. Noh, M. (1996) Simulation of Water Quality in Sewer Flow by Neural Networks, *Hydroinformatics '96, Proceedings of the 2nd International Conference on Hydro-informatics at Zurich Switzerland*, pp. 885-891.