

추계학적 비선형 모형을 이용한 달천의 실시간 수질예측

Real Time Water Quality Forecasting at Dalchun Using Nonlinear Stochastic Model

연인성^{1,*} · 조용진² · 김건홍³

Yeon, In-sung^{1,*} · Cho, Yong-jin² · Kim, Geon-heung³

1 충북대학교 토목공학과

2 충주대학교 환경공학과

3 인하대학교 환경토목공학부

(2005년 8월 11일 논문 접수: 2005년 12월 9일 최종 수정논문 채택)

Abstract

Considering pollution source is transferred by discharge, it is very important to analyze the correlation between discharge and water quality. And temperature also influent to the water quality. In this paper, it is used water quality data that was measured DO (Dissolved Oxygen), TOC (Total Organic Carbon), TN (Total Nitrogen), TP (Total Phosphorus) at Dalchun real time monitoring stations in Namhan river. These characteristics were analyzed with the water quality of rainy and nonrainy periods. Input data of the water quality forecasting models that they were constructed by neural network and neuro-fuzzy was chosen as the reasonable data, and water quality forecasting models were applied. LMNN (Levenberg-Marquardt Neural Network), MDNN (MoDular Neural Network), and ANFIS (Adaptive Neuro-Fuzzy Inference System) models have achieved the highest overall accuracy of TOC data. LMNN and MDNN model which are applied for DO, TN, TP forecasting shows better results than ANFIS. MDNN model shows the lowest estimation error when using daily time, which is qualitative data trained with quantitative data. If some data has periodical properties, it seems effective using qualitative data to forecast.

Key words: artificial intelligence, neural network, neuro-fuzzy, real time, water quality forecast

주제어: 인공지능, 신경망, 뉴로-퍼지, 실시간, 수질예측

1. 서 론

농업의 근대화 와 여가 선용에 따른 위락시설의 증

가로 인해 대도시뿐만 아니라 중 소도시 및 농어촌
지역에서도 각종 수질오염 물질의 배출량이 점진적
으로 늘어나고 있는 추세이다. 1998~2000년까지 한강
유역의 수질오염사고는 총 70회 중에 68.6%는 유류

*Corresponding author Tel: 82-43-261-2226, FAX: + 82-43-275-2377, E-mail: isyeon@chungbuk.ac.kr (Yeon, I.S.)

유출이며, 그 외에는 수환경 변화에 의한 것으로 조사되었다. 남한강에서는 유류유출이 17회, 수환경변화로 인한 오염사고가 3회 발생하였다. 유류유출 사고는 주로 차량사고와 유류저장탱크 관리 소홀 등에 의해서 발생되었으며, 수환경변화에 의한 사고는 강우유출수의 급격한 하천유입, 하상퇴적물 부상, 하천의 급격한 용존산소량의 고갈 및 각종 폐수의 불법 배출 등에 의해 발생하는 것으로 나타났다(환경관리공단, 2001). 이러한 변화에 대비하여 우리나라 하천들도 보다 진일보된 수질오염 관리 시스템에 접근할 필요가 있다.

현재 국내에서는 주요 하천에 수질자동측정망이 설치되어 운영되고 있다. 달천 수질 자동측정망지점은 충주시 살미면 토계리에 위치하고 있으며, 상수원 수질 감시 및 오염물질 배출업소 감시의 목적으로 운영되고 있다. 자동 측정되는 수질자료들을 보다 효과적으로 이용하기 위해서는 적절한 예측 모형을 통해 실시간 감시체제가 구축되어야 한다.

실시간으로 측정할 수 있는 자동측정 항목에는 한계가 있으며, 필요한 여러 가지 입력변수들을 모두 취득하는 것은 어렵기 때문에 예측 모형은 입력 자료를 최소화하여 적절한 시간 범위 내에서 예측을 위한 연산이 이루어질 수 있어야 한다(Tay 등, 2000). 따라서 연속적인 측정 및 실시간적 관리가 요구되는 수질자동측정망이나 취수장, 그리고 상습 침수구역에서 수질과 유량의 변화를 분석하고 예측하는 데 있어 추계학적 모형은 적합한 방법이라 할 수 있다. 하천 및 저수지의 수질예측에 추계학적 모형에 대한 연구들(류 등, 1998; 김, 2002; 정 등, 2003)이 진행되어 왔으며, 특히 입력과 출력만을 고려하여 구성이 용이하고 학습을 통해 변화에 잘 적응할 수 있는 인공지능 모형들은 유효적절한 모형이라 할 수 있다(안 등, 2000; 이 등, 2001; 김 등, 2003; 조 등, 2004; Mohammad 등, 2005).

이번 연구에서는 시계열 자료를 이용하여 추계학적 형태로 학습된 신경망 모형과 뉴로-퍼지 모형으로 실시간 수질예측 모형을 구성하고, 달천유역의 DO(Dissolved Oxygen), TOC(Total Organic Carbon), TN(Total Nitrogen), TP(Total Phosphorus)의 예측에 적용하였다. 먼저 5분단위로 측정된 달천 수질자동측정망 지점의 원시자료를 수집

하여 시간 대표값을 산출하였다. 강우시와 무강우시로 구분하여 유량, 수온과 수질인자들의 상관관계 및 주기적 성향 등 DO, TOC, TN, TP 자료의 특성을 분석하였다. 분석한 수질특성을 기준으로 수질예측 모형의 입력형태를 구성하고, 수질예측을 수행하였다. 수질예측 모형들은 은닉층 노드수와 학습회수를 변화하면서 학습하지 않은 검증 사상에 대한 오차를 통해 적절한 구조를 결정하도록 자동화되었다. 수질예측 결과를 분석하고, 예측 모형들에 대한 적용성을 검토하였다.

2. 적용 모형의 기본이론

2.1. 신경망 모형

신경망은 은닉층의 수에 따라 단층신경망과 다층신경망으로 구분된다. 입력층과 출력층만을 갖는 단층신경망은 출력값을 생성하는 활성화 함수(ϕ)로 계단함수를 사용하여 복잡한 과정을 생략하고 빠르게 계산할 수 있다. 입력층, 은닉층, 출력층을 갖는 다층신경망은 활성화 함수(ϕ)로 시그모이드(sigmoid) 함수와 이로부터 변형된 로그-시그모이드(log-sigmoid) 함수 등을 사용한다.

신경망 모형에서 가장 많이 이용되는 다층신경망 모형은 주로 BP(Back Propagation, 오차 역전파) 알고리즘을 이용한다. BP 알고리즘에서는 미분값을 사용하며, 학습하는 동안 목표값과 실제 출력값 사이의 계산된 오차 정보는 네트워크를 통하여 역방향으로 전달되고, 오차를 최소화시키는 방향으로 연결 가중치에 영향을 준다.

경사도 $\lambda = 1$ 인 단극성 시그모이드 함수를 전달함수로 사용하는 경우, 출력층의 가중합 u_j 로부터 출력층에서의 출력 Y 는 식 (1)과 같이 계산된다. 식 (2)에 나타난 출력층의 오차신호 δ_j 는 식 (3)의 미분결과로부터 식 (4)와 같이 목표값 T 와 출력값 Y 로 표현되며, 식 (5)와 같이 은닉층 출력 H 와 연산되어 연결강도 조정에 영향을 미친다.

$$Y = \phi(u_j) = \frac{1}{1 + \exp(-u_j)} \quad (1)$$

$$\delta_j = (T - Y) \phi'(u_j) \quad (2)$$

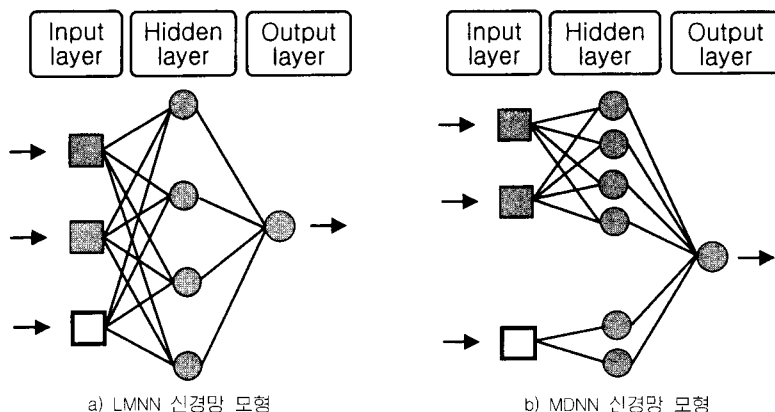


Fig. 1. 신경망 모형의 구조.

$$\varphi(u_j) = \frac{\exp(-u_j)}{[1 + \exp(-u_j)]^2} \quad (3)$$

$$= \frac{1}{1 + \exp(-u_j)} \times \frac{1 + \exp(-u_j) - 1}{1 + \exp(-u_j)}$$

$$= \varphi(u_j) [1 - \varphi(u_j)]$$

$$= Y(1 - Y)$$

$$\delta_j = (T - Y)Y(1 - Y) \quad (4)$$

$$\Delta W = \delta_j H \quad (5)$$

여기서 ΔW 는 은닉층과 출력층간의 연결강도 변화량이다. 출력층과 은닉층에서 계속적인 연결강도의 갱신을 통해 모형을 최적화하며, 오차가 지정한 범위 내로 만족하거나 학습회수에 도달하면 학습을 종료하고, 학습을 마친 연결강도를 사용하여 예측을 수행하게 된다.

일반적인 BP 알고리즘에 비해 우수성이 여러 연구를 통해 입증된 Levenberg-Marquardt 방법은 연결강도를 조정할 때 식 (6)과 같은 Levenberg-Marquardt update rule을 사용하여 반복 계산되는 η 의 크기를 효율적으로 제어하는 방법이다.

$$\Delta W = (J^T J + \eta I)^{-1} J^T e \quad (6)$$

여기서 J 는 각 연결강도에 대한 오차 미분의 Jacobian matrix, η 는 적응학습율, e 는 오차벡터이다.

일반적인 신경망 모형의 연결강도는 보통 모든 뉴

런에 연결되어져 연산되나 Modular 방법은 전체 네트워크를 부분적인 네트워크(sub-network)로 분류할 수 있다. 성질이 다른 입력변수를 구분하고 해당 은닉층에만 영향을 미치도록 구성되어 있어 입력 자료별로 나누어서 학습을 진행시키므로 학습 시간을 단축시킬 수 있을 뿐만 아니라 정상적인 자료와 정성적인 자료의 특성을 살려 동시에 학습시킬 수 있다. 시간 변화에 많은 영향을 받는 수질 자료의 특성을 고려하여 정성적 자료인 시간항목을 독립적으로 연산하고 최종 출력값에만 영향을 미치도록 할 수 있다.

수질예측을 위해 사용된 신경망 모형으로 은닉층을 분리하지 않은 LMNN(Levenberg-Marquardt Neural Network) 모형, 은닉층을 분리한 MDNN(MoDular Neural Network) 모형의 구조를 Fig. 1에 나타내었다.

2.2. ANFIS 모형

ANFIS(Adaptive Neuro-Fuzzy Inference Systems)는 신경회로망과 퍼지이론이 결합된 형태로서 신경회로망의 구조와 학습능력을 이용하여 제어환경으로부터 얻은 입출력 정보로부터 소속함수와 제어규칙을 제어 대상에 맞게 자동 조절하는 방식의 모형이다(Jang, 1993). ANFIS의 기본적인 학습방법은 각 노드의 출력에 대한 제공된 오차의 미분 즉, 오차 신호를 계산하는 역전파 알고리즘에 의해 후방향 경로에서 전체부 매개변수를 조정한다. 소속함수의 모양이 전체부 매개변수에 의해 정의되어지기 때문에 제공오차 E 는 결론부 매개변수가 고정된다면 소속함수의 폭과 중심

을 계산하는 $\{a_i, c_i\}$ 의 매개변수로 결정되어진다. E의 값을 감소시키기 위해서는 전체부 매개변수에 대해서 다음과 같이 최대 경사법(steepest descent method)에 의해 반복적으로 이루어진다.

$$a_i(t+1) = a_i(t) - \eta \frac{\partial E}{\partial a_i} \quad (7)$$

$$c_i(t+1) = c_i(t) - \eta \frac{\partial E}{\partial c_i} \quad (8)$$

$$\eta = \frac{k}{\sqrt{\sum_a \left(\frac{\partial E}{\partial a}\right)^2}} \quad (9)$$

여기서 a 는 전체부 매개변수이고, k 는 매개변수 공간에서 경사벡터의 이동한 거리로 수렴 속도 변화에 영향을 미친다. 전체부 매개변수가 주어질 때 전체적인 출력 f 는 결론부 매개변수인 $\{p_i, q_i, r_i\}$ 의 선형적인 조합에 의해 식 (10)과 같이 표현되어질 수 있다.

$$\begin{aligned} f &= \bar{w}_1 f_1 + \bar{w}_2 f_2 \quad (10) \\ &= (\bar{w}_1 x) p_1 + (\bar{w}_1 y) q_1 + (\bar{w}_1) r_1 + (\bar{w}_2 x) p_2 + (\bar{w}_2 y) q_2 \\ &\quad + (\bar{w}_2) r_2 \end{aligned}$$

3. 대상유역 및 자료의 전처리

3.1. 대상유역

대상유역인 달천유역은 남한강의 주요 지류중의 하나이며, 유역면적 1,614.36km²이다. 유로연장은 124.39km, 유역의 평균 폭은 12.85km이다. 달천 유역을 Fig. 2에 나타내었다.

달천 유역에는 달천 본류로 유입되는 음성천 등 17개의 제 1지류가 있으며, 미원천 등 15개의 제 2지류로 형성되어 있다. 달천은 제 1지류인 쌍천과의 합류점을 지나면서 본류의 하폭은 최소 90m로부터 최대 340m 정도이고, 하상은 굽은 자갈로 형성되어 있다.

달천 수질자동측정망은 충주시 살미면 토계리에 위치하고 있으며, 토계취수장 취수원수에 대한 수질 감시와 음성공단 폐수로 인한 오염사고에 대비할 목적으로 운영되고 있다. 달천 수위관측소는 수질자동측정망 직하류인 살미면 향신리에 위치해 있다.

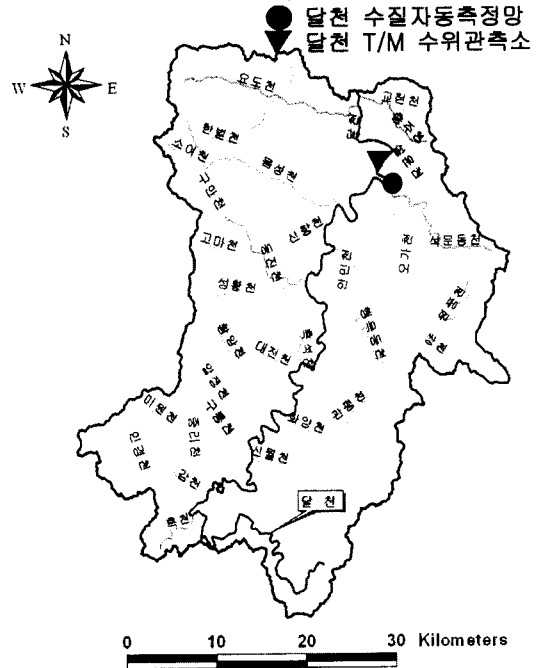


Fig. 2. 달천 유역도.

3.2. 자료의 전처리

연구를 위해 사용된 달천 수질자동측정망의 수온, DO, TOC, TN, TP 자료는 2003년의 1월, 4월, 7월의 자료로서 5분단위로 수집되었다. 취득 가능한 유량 자료와 연계하여 모의 및 예측을 수행하기 위해서는 자료들간의 시간간격을 일치시켜야 한다. 수위관측소로부터 동일기간에 산정된 유량 자료는 시단위 자료이므로 수질자동측정망에서 매 5분마다 취득된 자료들을 시단위 대표값으로 산출하는 과정을 거쳤다. 시단위 대표값이 평균값으로 산출될 경우 순간적인 기기의 오작동에 의한 부정확한 자료를 포함할 수 있으므로 안정화된 대표값을 산출하기 위해서 중앙(median)값을 추출한 후 이동평균(moving average)법을 이용하여 평활화 과정이 수행되었다. Fig. 3에 수집된 자료중 DO의 중앙값 추출결과를 나타내었다.

중앙값 추출을 통해 전반적으로 양질의 시단위 자료를 취득하였으나, 부분적으로 불안정한 구간이 존재하여 2차적인 안정화 과정이 요구되었다. 5분단위 원시자료가 1시간 동안 모두 소멸된 구간에서는 전후 자료를 통해 보간되었으며, 취득된 시단위 대표값이 진동하는 구간에서는 자기중심 이동평균법을 통해 평

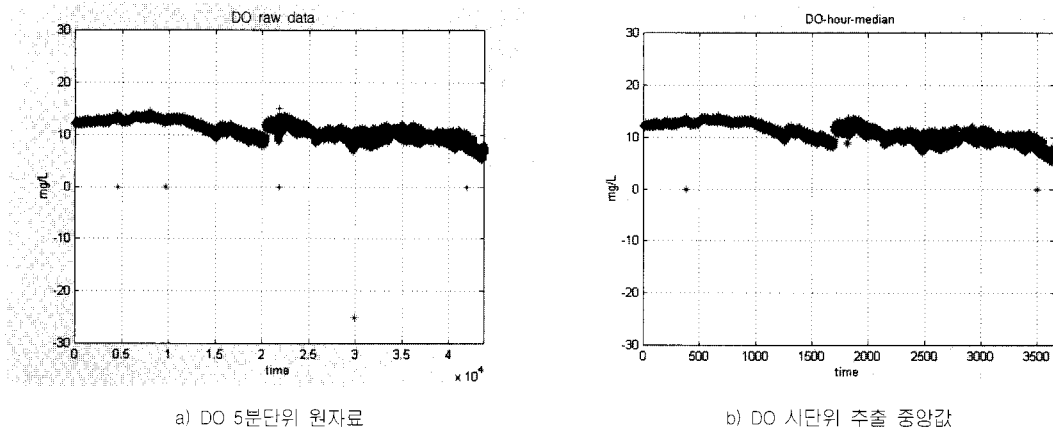


Fig. 3. 5분단위 원자료의 중앙값 추출 결과.

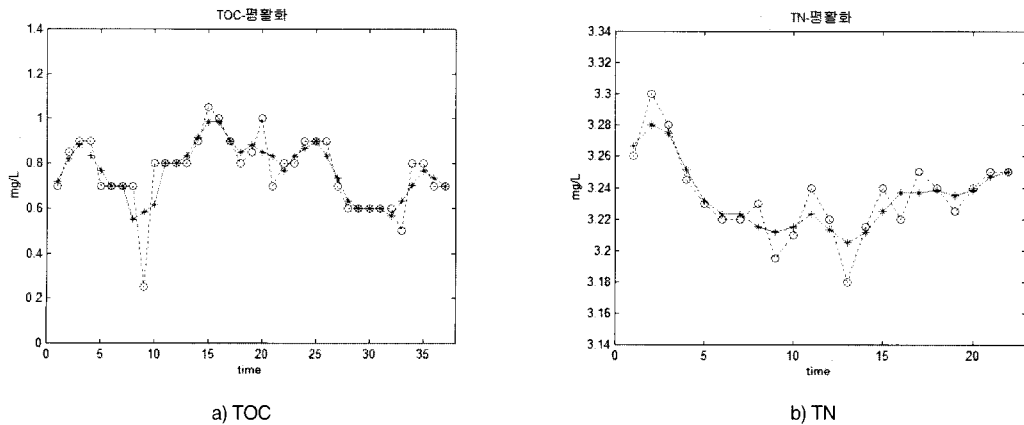


Fig. 4. 자료의 평활화 결과(...○...: 중앙값, -*-: 평활화).

활화 과정이 수행되었다. Fig. 4에 수질 자료 중 TOC와 TN의 평활화 결과를 나타내었다.

3.3. 자료 특성 분석

전처리 과정을 통해 취득된 자료들을 강우의 유무에 따라, 계절에 따라, 시간대별 특성에 따라 분석하였다. 분석을 통해 수질예측 모형의 구조를 결정하고자 하였다. 강우의 유무에 따라 강우시와 무강우시로 구분한 자료특성을 Table 1에 나타내었다. 달천 지점에서 DO 농도의 평균값은 강우시에 5.5mg/L, 무강우시에 10.0mg/L로 무강우시의 평균값이 큰 것으로 나타났다. 강우시 DO 농도는 수질의 탁도와 연계되어 작아졌으며, 1월 자료의 대부분이 무강우시에 해당되었기 때문에 수온의 영향으로 인해 무강우시의

표준편차가 더 큰 것으로 판단된다. 강우시의 TOC, TN 농도는 무강우시의 자료에 비해 최소값, 평균값, 최대값이 크게 나타났으며, 표준편차는 미소한 차이를 보였다. TN 농도의 평균값은 무강우시에 크게 나타났다. 최대값은 강우시에 큰 것으로 나타났다.

하천의 수질은 시기에 따라 다른 특성을 갖는다. 각 지점에서 측정된 수온, DO, TOC를 중심으로 1월, 4월, 7월의 계절별 특성을 분석하였다. 달천 지점의 월별 DO 평균값은 1월에 13.7mg/L, 4월에 9.6mg/L, 7월에 4.8mg/L로 여름철인 7월에 가장 낮았으며 월별 TOC 평균값은 1월에 0.7mg/L, 4월에 1.1mg/L, 7월에 1.4mg/L로 7월에 가장 큰 값을 보였다. 달천 지점의 수위와 유량의 평균값은 7월에 각각 1.05m와 144.16cms로 가장 컸으며 DO, TOC, 수위

Table 1. 각 항목별 적용 자료의 특성

항목	강우시						무강우시					
	수온 (°C)	DO (mg/L)	TOC (mg/L)	TN (mg/L)	TP (mg/L)	유량 (cms)	수온 (mg/L)	DO (mg/L)	TOC (mg/L)	TN (mg/L)	TP (mg/L)	유량 (cms)
최소값	12.80	2.70	0.80	1.33	0.04	6.42	0.50	3.90	0.40	1.82	0.01	0.82
평균값	20.00	5.50	1.40	2.27	0.16	145.52	12.50	10.00	0.90	2.69	0.10	9.36
최대값	24.60	10.80	2.20	3.76	0.54	1843.02	25.80	14.40	1.70	3.25	0.29	42.12
표준편차	2.90	2.00	0.25	0.74	0.10	243.69	7.60	2.90	0.24	0.35	0.08	8.27

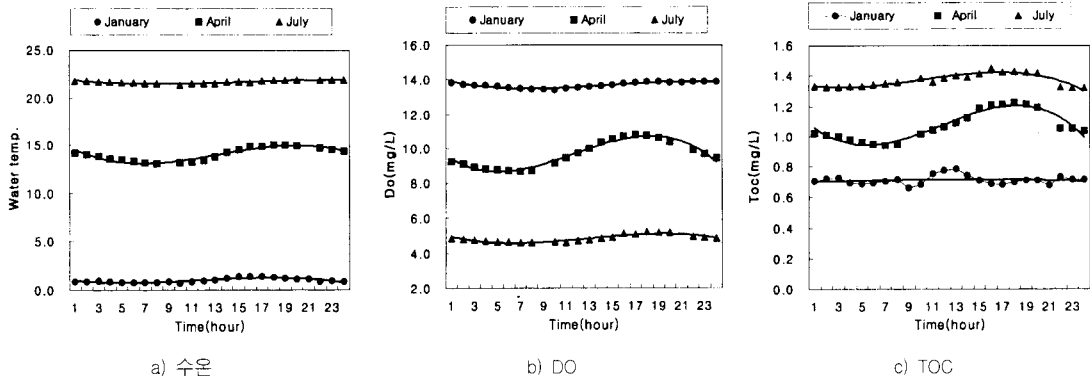


Fig. 5. 달천 자료의 시간대별 평균값.

및 유량의 표준편차 또한 7월에 가장 큰 것으로 분석되었다. 1월, 4월, 7월의 각 시간별 자료특성을 분석하였으며, 시간대별 평균값을 Fig. 5에 나타내었다.

수온, DO, TOC는 계절별로 큰 차이를 보여 계절적인 특성을 가지는 것으로 나타났다. 특히 봄철에는 수질이 일정한 변화 경향을 보였다. 자정부터 서서히 감소하여 일출시간을 기점으로 증가하고, 16~19시에서 다시 감소하였다. 수온과 유량의 변화는 수질에 많은 영향을 미치는 것으로 분석되었다.

4. 적용 및 결과

4.1. 모형의 구성

수질변화에 주요한 영향을 미치는 인자가 수온과 유량임을 감안하여 각 항목별로 교차상관 분석을 수행하였으며, 결과에 따라 각 항목별 입력자료를 Table 2와 같이 선정하였다. 강우시에는 유량과의 상관성이 높았으며, 무강우시에는 수온과의 상관성이 대체로 높았으나 항목에 따라 상이한 결과가 분석되어 교차상관 분석결과를 기준으로 하였다. 수온은

DO와 0.90, TOC와 0.77의 상관계수를 보였고 유량은 TOC, TN, TP와 0.60 내외의 상관계수를 보였다. LMNN 모형과 ANFIS 모형은 2개의 입력값을 사용하였으며 t시간의 수질자료와 함께 t시간의 유량 혹은 수온 자료를 입력값으로 사용하여 t+1시간에서의 수질을 출력하도록 하였다. 은닉층을 분리한 MDNN 모형은 정성적 자료인 시간개념을 추가하여 t+1시간에서의 수질을 출력하도록 하였다. 정성적 자료로서 강우시에는 강우발생후 경과시간을 이용하였으며, 무강우시에는 일조 및 일몰의 영향을 고려하여 수질 및 유량이 측정된 시간을 이용하였다. 추계학적 모형은 예측하고자하는 수질자료들의 변화성향도 고려되며, 연구에 사용된 자료들의 자기상관계수는 lag 1인 경우가 가장 높았다. 따라서 사용된 자료가 시간위 자료이므로 예측 간격은 최소 단위인 1시간 예측으로 결정하였다. 달천 수질자동측정망은 취수 원수의 실시간 감시목적상 5분 단위로 측정되므로 유량 및 기타 현장자료의 취득이 가능하다면 더 짧은 간격의 예측이 보다 유효할 것으로 판단된다.

인공지능 모형의 구성을 위해서는 학습절차를 거

Table 2. 수질예측 모형의 인출력 구성

항목	모형	입력		출력
		강우시	무강우시	
DO	LMNN	DO(t), 수온(t)	DO(t), 수온(t)	DO(t+1)
	ANFIS	DO(t), 수온(t)	DO(t), 수온(t)	
	MDNN	DO(t), 수온(t), 강우진행시간	DO(t), 수온(t), 일일진행시간	
TOC	LMNN	TOC(t), 유량(t)	TOC(t), 수온(t)	TOC(t+1)
	ANFIS	TOC(t), 유량(t)	TOC(t), 수온(t)	
	MDNN	TOC(t), 유량(t), 강우진행시간	TOC(t), 수온(t), 일일진행시간	
TN	LMNN	TN(t), 유량(t)	TN(t), 유량(t)	TN(t+1)
	ANFIS	TN(t), 유량(t)	TN(t), 유량(t)	
	MDNN	TN(t), 유량(t), 강우진행시간	TN(t), 유량(t), 일일진행시간	
TP	LMNN	TP(t), 유량(t)	TP(t), 유량(t)	TP(t+1)
	ANFIS	TP(t), 유량(t)	TP(t), 유량(t)	
	MDNN	TP(t), 유량(t), 강우진행시간	TP(t), 유량(t), 일일진행시간	

쳐야 하며, 적절한 은닉층의 노드수와 학습회수 등을 조정해야 하는 반복적인 과정이 요구된다. 자료의 형태와 특성에 따라 적정 노드수는 달라질 수 있으며, 선형 관계와 달리 비선형적인 관계를 갖는 자연현상 및 수문현상을 모의할 때에 과도한 학습은 학습하지 않은 사상에 대한 모형의 검증시에 큰 오차를 유발한다. 그러나 지금까지 노드수와 학습회수에 대해 연구된 바는 있으나 확정된 결론을 얻지는 못하였다.

이번 연구에서는 은닉층의 노드수와 학습회수를 변화해가며 학습하지 않은 검증사상에 적용하여 최소의 오차를 갖는 구조를 찾도록 하였고, 수동적인 반복과정을 피하기 위해 검증절차를 자동화하였다. 인공지능 모형의 자동화된 매개변수 산정과정을 Fig. 6에 나타내었다.

실시간으로 측정되는 유량과 수온 및 수질 자료로부터 수질예측 모형의 입력자료가 선정되었으며, 충분한 학습을 거쳐 입력자료와 출력자료간의 관계를 해석할 수 있는 수질예측 모형이 완성되었다. 모형의 검증시에 학습회수를 증가할수록 학습에 의한 검증 오차는 작아진다. 그러나 학습에 의한 검증오차를 기준으로 적정모형을 선정한다면 예측시에 큰 오류를 범할 수 있다. 학습하지 않은 자료에 대해서 예측을 수행할 경우, 모형의 예측에 따른 검증오차는 학습에 의한 검증오차와 반드시 비례하지 않기 때문이다. 입력층, 은닉층, 출력층에 대한 구성 형태에 따라서 모형의 적정한 학습회수는 다를 것이며, 선정된 입력자

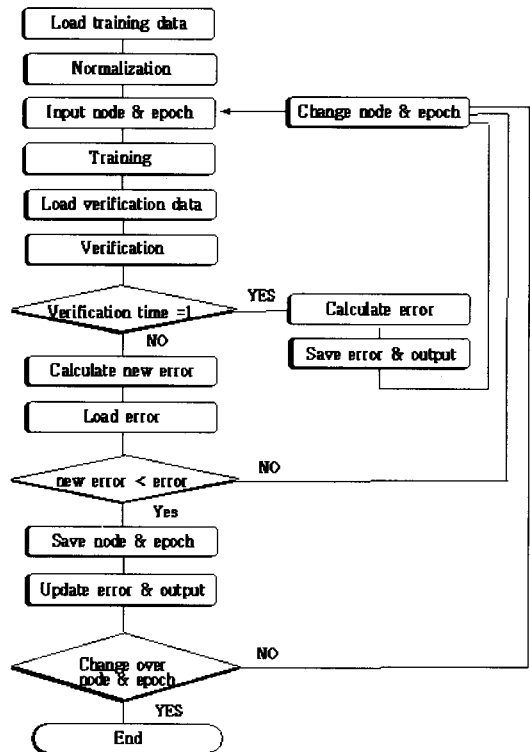


Fig. 6. 인공지능 모형의 자동화된 매개변수 산정 과정.

료 및 지점별 수질변화 양상에 따라서 모형의 적정한 구성 형태 또한 다를 것으로 판단된다. 따라서 연구 진행을 위해 모형의 선정에는 검증오차를 기준으로 하였으며, 검증 오차를 최소화하는 방향으로 적정 은

Table 3. 선정된 수질항목별 최적모형의 구조

구분	항목	LMNN	MDNN	ANFIS
강우시	DO	2-4-1 (1000)	3-[4,1]-[4]-1 (500)	2-4-1 (5)
	TOC	2-6-1 (2500)	3-[2,2]-[2]-1 (1000)	2-2-1 (5)
	TN	2-2-1 (4000)	3-[2,2]-[4]-1 (1000)	2-2-1 (500)
	TP	2-2-1 (5000)	3-[1,1]-[1]-1 (500)	2-2-1 (5)
무강우시	DO	2-6-1 (3000)	3-[4,2]-[1]-1 (5000)	2-2-1 (40)
	TOC	2-2-1 (500)	3-[2,1]-[4]-1 (500)	2-2-1 (50)
	TN	2-4-1 (2000)	3-[4,1]-[4]-1 (2000)	2-2-1 (10)
	TP	2-4-1 (4000)	3-[4,1]-[4]-1 (500)	2-2-1 (500)

주) LMNN:Inputnode-Hiddennode-Outputnode(epoch)

MDNN:Inputnode-[Hiddenlayer1node-Hiddenlayer2node]
-[Outputlayer]-Outputnode(epoch)

ANFIS:Inputnode-Hiddennode-Outputnode(epoch)

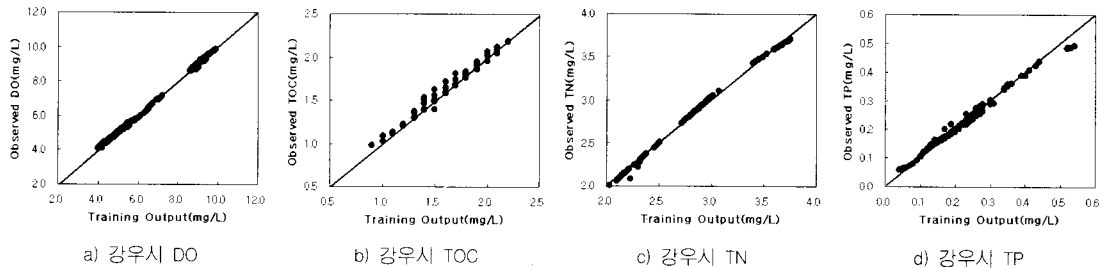


Fig. 7. 학습 결과.

Table 4. 각 모형별 강우시 수질예측 결과

모형	강우시							
	DO		TOC		TN		TP	
	RMSE(mg/L)	CC	RMSE(mg/L)	CC	RMSE(mg/L)	CC	RMSE(mg/L)	CC
LMNN	0.201	0.979	0.100	0.910	0.088	0.978	0.027	0.886
MDNN	0.277	0.959	0.090	0.930	0.073	0.985	0.027	0.886
ANFIS	0.900	0.546	0.090	0.930	0.152	0.973	0.070	0.836

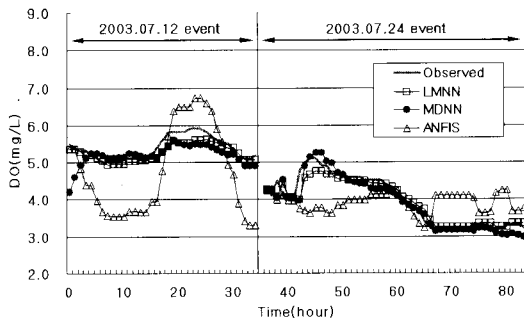
닉층의 형태와 학습회수를 찾도록 반복과정을 자동화 하였다. 모형의 자동화를 통해 지점별, 항목별로 가장 적합한 결과를 보인 수질예측 모형을 선정하였으며, 선정된 모형의 구성 및 연결강도를 Table 3에 나타내었다.

구성된 모형들은 학습회수와 은닉층의 노드수를 변화해가며, 스스로 적절한 구조를 찾도록 학습되었다. Fig. 7에 적용한 모형 중 MDNN 모형의 학습 결과를 나타내었다. 강우시와 무강우시의 시단위 수질 자료에 대한 변화양상을 학습하였으며, 수질변화에

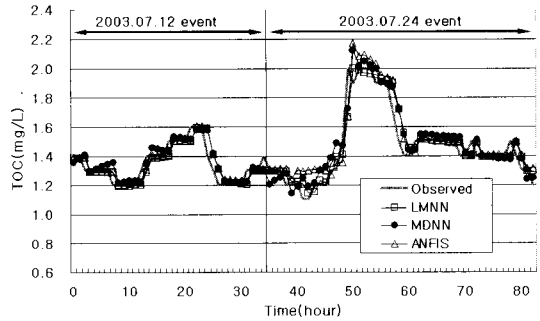
대한 충분한 학습이 이루어졌음을 알 수 있다.

학습된 신경망 모형들과 뉴로-퍼지 모형을 적용하여 수질 예측된 결과를 Table 4와 Fig. 8에 나타내었다.

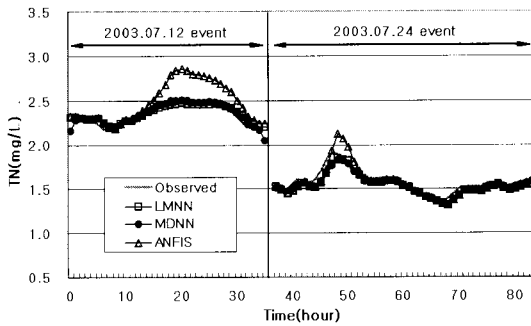
ANFIS 모형은 강우시 TOC 예측에 비교적 좋은 결과를 보였으나 강우시 DO, TN, TP 예측에서는 LMNN, MDNN 모형에 비해 적용성이 크게 떨어지는 것으로 나타났다. LMNN, MDNN의 두 신경망 모형은 모든 항목에서 안정적인 예측결과를 보였으며, 정성적 자료인 강우발생시간을 이용한 MDNN



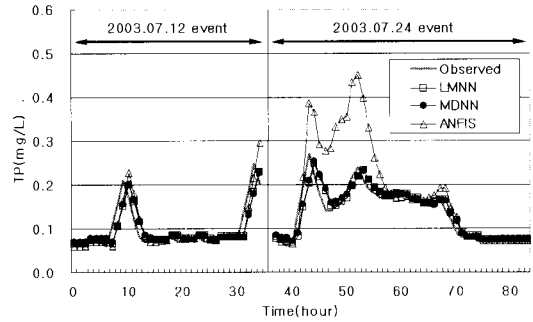
a) 강우시 DO 예측 결과



b) 강우시 TOC 예측 결과



c) 강우시 TN 예측 결과



d) 강우시 TP 예측 결과

Fig. 8. 강우시 수질예측 결과.

Table 5. 각 모형별 무강우시 수질예측 결과

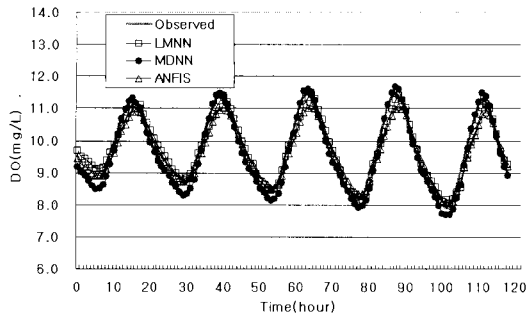
모형	무강우시							
	DO		TOC		TN		TP	
	RMSE(mg/L)	CC	RMSE(mg/L)	CC	RMSE(mg/L)	CC	RMSE(mg/L)	CC
LMNN	0.274	0.968	0.057	0.896	0.017	0.995	0.031	0.907
MDNN	0.213	0.991	0.048	0.932	0.016	0.996	0.026	0.938
ANFIS	0.316	0.966	0.057	0.896	0.024	0.995	0.030	0.912

모형이 비교적 양호한 적용성을 보였다. 신경망 모형의 예측결과에 대한 상관계수는 DO, TOC, TN에 비해 TP가 다소 낮은 것으로 나타났다.

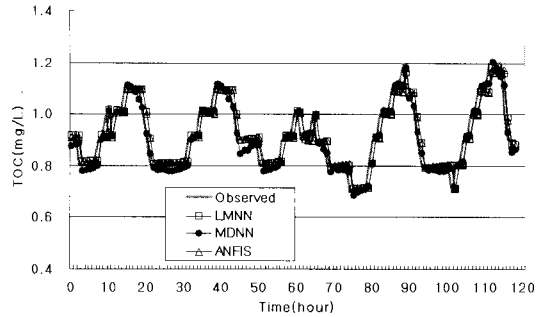
무강우시의 수질예측 결과를 Table 5와 Fig. 9에 나타내었다. 무강우시 DO 예측과 TN 예측에 적용한 LMNN, MDNN, ANFIS 모형 모두 실측치에 근사하였으며 높은 상관성을 보였다. 무강우시 TOC와 TP 예측에서는 세 모형이 모두 실측치에 근사하였으나, DO 및 TN의 예측결과에 비해 상관계수는 낮은 것으로 나타났다. 무강우시 수질예측에 적용한 모형

들중 MDNN 모형이 가장 낮은 오차와 높은 상관계수를 보였다.

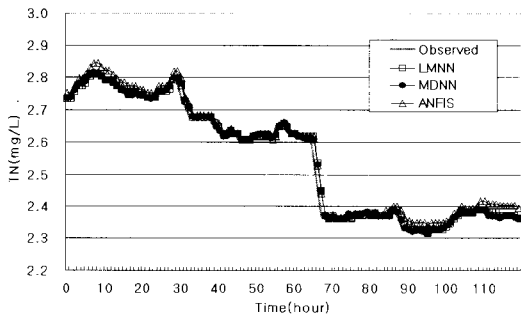
유량 및 수질 자료가 축적되어 있는 하천에서 신경망 모형은 매우 안정적인 적용성을 보였다. 또한 운영 및 계산시간이 매우 짧아 빠르게 대처해야 하는 수질오염사고에 대비하여 적합한 모형이 될 수 있다고 판단된다. 유량 및 수질 자료가 부족하고 양상이 불규칙적으로 급변하는 지역에서는 강우발생시간과 같은 정성적 자료의 활용이 모형의 적용성 향상에 도움이 될 것으로 판단된다.



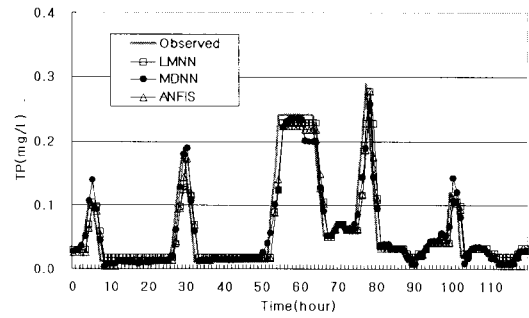
a) 무강우시 DO 예측 결과



b) 무강우시 TOC 예측 결과



c) 무강우시 TN 예측 결과



d) 무강우시 TP 예측 결과

Fig. 9. 무강우시 수질예측 결과.

5. 결 론

수질자동측정망이 운영되고 있는 달천 지점의 실시간 수질 자료를 이용하여 특성을 분석하고, 신경망 모형과 뉴로-퍼지 모형을 구성하여 수질예측에 대한 적용성을 검토하였으며, 다음과 같은 결론을 얻었다.

1. 수질자료를 강우의 유무에 따라 분석한 결과, 강우시에 측정된 DO 농도는 무강우시에 측정된 자료보다 최소값, 평균값, 최대값이 낮았으며, 강우시에 측정된 TOC, TP 자료는 무강우시 측정된 자료에 비해 최소값, 평균값, 최대값이 크게 나타났다. 계절별로는 여름철의 수질이 가장 악화되었으며, 강우에 따른 유량의 변화가 수질변화에 크게 영향을 미치는 것으로 판단된다. 봄철에는 또한 24시간을 기준으로 주기적인 수질변화 성향이 나타났다.

2. 달천지점에서 강우시 수질예측을 위해 적용한 LMNN, MDNN의 두 신경망 모형은 안정적인 적용

성을 보였으며, 정성적 자료인 강우발생시간을 활용한 모형이 비교적 우수한 예측 결과를 보였다. ANFIS 모형은 달천 지점의 TOC 예측에 비교적 양호하였으나, 이외의 예측에서는 신경망 모형에 비해 큰 오차를 보였다. 무강우시 DO, TOC, TN, TP 예측에 적용한 LMNN, MDNN, ANFIS 모형은 모두 실측치에 근사하고 작은 오차를 보였으며, MDNN 모형이 타 모형에 비해 비교적 좋은 결과를 보였다.

은닉층을 분리한 MDNN 모형의 적용 결과가 비교적 우수한 것으로 나타남으로써 수질 예측시에는 정량적 자료와 함께 정성적 자료의 활용이 효과적인 것으로 판단된다. 수질자동측정망으로부터 취득되는 자료가 더욱 축적되어 안정화된다면 수질예측 능력은 향상될 것이며, 취수시설이 위치한 곳이나 수질변화에 민감한 지점에서 실시간 수질예측 모형은 유용하게 사용될 수 있을 것으로 판단된다.

참고문헌

- 김만식, 한재석 (2003) 인공지능기법을 이용한 수질모형의 최적 매개변수 추정연구, *환경관리학회지*, **9**(1), pp. 1-9.
- 김주식, 서동일, 이은형, 이준기 (2002) 금강수계 수질자동 측정망 추가설치를 위한 우선순위 분석, *대한상하수도학회 공동춘계 학술발표회 논문집*, pp. 61-63.
- 김주환 (2002) 갈수기 댐 하류 수질예측을 위한 신경망 모형의 적용, *한국수처리기술연구회지*, **10**(4), pp. 51-60.
- 류병로, 한양수 (1998) ARIMA 모형에 의한 하천수질 예측, *한국환경과학회지*, **7**(4), pp. 433-440.
- 안상진, 연구방, 연인성 (2000) 신경망 모형을 이용한 수질 예측과 시스템 구축, *대한상하수도학회 공동춘계 학술발표회 논문집*, pp. 249-252.
- 이경훈, 김진모, 문병석 (2001) BP 알고리즘을 이용한 호소 수 수질예측, *대한상하수도학회 공동추계 학술발표회 논문집*, pp. 203-206.
- 정세용, 강병수, 조주영 (2003) 일별 암모니아성 질소농도 예측을 위한 다중회귀모형 개발, *대한환경공학회 추계 학술연구발표회 논문집*, pp. 136-144.
- 조용진, 연인성, 이재관 (2004) 실시간 수질예측을 위한 신경망 모형의 적용, *한국물환경학회지*, **20**(4), pp. 321-326.
- 환경관리공단 (2001) 팔당호등 한강수계 수질자동측정망 확대설치타당성조사 결과보고서, pp. 44-46.
- Jang, J.S.R. (1993) ANFIS: Adaptive-Network based Fuzzy Inference System, *IEEE transactions on Systems, Man, and Cybernetics*, **23**(3), pp. 665-685.
- Simon Haykin (1999) *Neural Networks*. 2nd Ed. Prentice Hall International, Inc.
- Tay, J.H. and Zhang, X. (2000) A Fast Prediction Neural Fuzzy Model for High-rage Anaerobic Wastewater Treatment Systems, *Water Resources Research*, **34**(11), pp. 2849-2860.
- Mohammad N.A. and Jagath, J.K. (2005) Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data, *Environmental modelling & software*, **20**(7), pp. 851-871.