

인공신경망기법을 이용한 하천수질인자의 예측모델링 - BOD와 DO를 중심으로 -

조 현 경
영남이공대학 토목과
(2000년 6월 19일 접수)

Predictive Modeling of River Water Quality Factors Using Artificial Neural Network Technique - Focusing on BOD and DO -

Hyeon-Kyeong Cho

Dept. of Civil Engineering, Yeungnam College of Science & Technology, Taegu 705-037, Korea
(Manuscript received 19 June, 2000)

This study aims at the development of the model for a forecasting of water quality in river basins using artificial neural network technique. Water quality by Artificial Neural Network Model forecasted and compared with observed values at the Sangju 1 and Dalsung stations in Nakdong river basin. For it, a multi-layer neural network was constructed to forecast river water quality. The neural network learns continuous-valued input and output data. Input data was selected as BOD, DO, discharge and precipitation. As a result, it showed that method III of three methods was suitable more than other methods by statistical test(ME, MSE, Bias and VER). Therefore, it showed that Artificial Neural Network Model was suitable for forecasting river water quality.

Key Words : river water quality, Artificial Neural Network Model

1. 서 론

최근들어 산업의 발달에 따라 인간의 생활수준은 크게 향상되었으나 환경오염은 날로 심화되고 있다. 특히 산업폐수 및 생활하수 등에 의한 상수원 오염의 가능성이 높아지면서 시민들 가운데 대다수가 가장 시급하게 추진해야 할 환경정책으로 수질문제를 들고 있으며, 수도물을 식수로 사용하기를 꺼리는 등 수질에 대한 불신감이 심각한 상태이다. 또한 하천은 유역사가지의 개발 확산과 제방축조 등으로 자정능력을 잃고 있어 수질오염을 가속화시키고 있다.

따라서 많은 자치단체가 환경정책 중에서 수질오염 방지 및 개선에 많은 예산을 투입하고 있다. 이에 따라 1970년대 이후 급속한 산업화 및 도시화와 더불어 급속히 높아지던 낙동강의 BOD 수치가 현재 한자리 수로 낮아졌으며, 이는 지금까지의 수질개선 노력이 큰 성과를 거두고 있다는 것을 의미한다. 그러나 현재의 낙동강 수계의 수질은 아직도 여전히 많은 문제점을 안고 있는 실정이다. 현재 국내 상수원 수질기준 항목은 45개에 불과한 반면 미국은 115개, 일본은 85개의 측정항목을 두

고 있다. 다시 말해서 국내에서 관리하지 않고있는 오염물질이 문제를 일으킬 수도 있는 것이다.

특히 영남지역의 젖줄인 낙동강은 하천을 따라 많은 도시들이 들어서 있어 수질오염의 가능성이 어느 지역보다 높다고 할 수 있다. 따라서 하천수질의 관리를 위한 계획을 수립하기 위해서는 정확한 수질예측이 필수적으로 선행되어야 한다. 그러나 과거의 장기적인 수질자료가 부족한 낙동강을 포함한 우리나라 사정으로 인하여 정확한 수질예측에 많은 문제점을 안고 있다.

그런데 기존에 사용되어 온 수질예측 방법들은 모형의 불확실성, 입력자료의 불확실성 및 매개변수의 불확실성 등에 의해 계산결과와 실측자료가 일치하지 않는 경향이 많았다. 그래서 결과에 영향을 미치는 불확실성을 제거하기 위한 노력의 일환으로 최근 공학분야에서 많이 이용되고 있는 최신 정보이론인 인공신경망기법을 수질예측에 적용하려는 시도가 있어 왔다.

지금까지 수질예측에 관한 국외의 연구문헌들을 살펴보면, Huck과 Farquhar¹⁾이 Box-Jenkins 방법을 이용하여 하천수질을 모의하였으며, Mehta 등²⁾이 Passaic강 수질의 추계학적 변화를 연구하였고, Thomann³⁾이 수질자료의 시계열분석을 실시한 바 있다. 또한 Chadderton 등⁴⁾이 DO모형의 불확정성 해석을 수행하였으며, Donigian

* 이 논문은 1999학년도 영남이공대학 연구조성비 지원에 의한 것임.

과 Huber⁵⁾가 도시지역 및 비도시지역에서의 비점원 오염원에 대한 수질을 모의하였고, Wanielista 등⁶⁾이 수질에 대한 비점원 오염원의 영향에 관해 발표한 바 있다. 국내에서는 권혁윤⁷⁾이 하천 유량변동에 따른 수질변화에 관한 연구를 수행하였으며, 유희정⁸⁾이 수문인자에 의한 하천수질변화에 관한 연구를 발표하였고, 신현석과 윤용남⁹⁾이 SWMM 모형을 이용하여 도시지역의 유출 및 NPS 오염물질의 배출을 모의한 바 있다. 그리고 이상수¹⁰⁾는 금강에서 ARIMA 모형을 이용하여 수질을 예측하였으며, 양우석¹¹⁾은 하천유량 변화에 따른 수질오염도 분석을 실시하였고, 백경원 등¹²⁾은 QUAL2E 모형을 이용하여 한강하류부에서의 수질해석을 수행한 바 있다. 또한 이상도¹³⁾는 중소하천의 유량과 수질변동 특성에 관한 연구를 수행하였으며, 김춘수¹⁴⁾는 남강댐 유역의 하천수질을 예측한 바 있다.

한편 인공신경망기법에 의한 수질예측에 관한 국외의 연구문헌을 살펴보면, Maier와 Dandy¹⁵⁾가 신경망이론에 의해 오스트레일리아 남부에 있는 Murray강의 수질을 예측한 바 있다. 국내에서는 강두선 등¹⁶⁾이 신경망 모형을 이용하여 한강유역에서 장기하천수질을 예측하였고, 신현석 등¹⁷⁾이 신경망을 이용하여 도시지역 유출 및 비점원 오염물 배출을 모의한 바 있다.

본 연구에서는 낙동강유역의 하천수질 관리를 위한 하천유지용수 공급계획, 하수처리장 신설 및 증설계획과 같은 중·장기적인 수질관리 대책을 수립함에 있어서 사용될 수 있는 인공신경망기법을 이용한 수질예측 모형을 개발하고, 모형의 적용 가능성을 제시하는데 그 목적을 두었다.

2. 인공신경망기법의 기본이론

2.1. 자연신경망의 개요

인간의 두뇌에는 약 140억 개나 되는 뉴런(Neuron, 신경세포)이 있고, 각 뉴런은 다음 Fig. 1에 나타난 바와 같이 본체인 세포체 부분과 수상돌기 부분 그리고 축삭 부분으로 구성되어 있다. 이들이 기억이나 학습 더 나아가 감정이나 마음이라는 고도의 정신활동을 담당하고 있다.

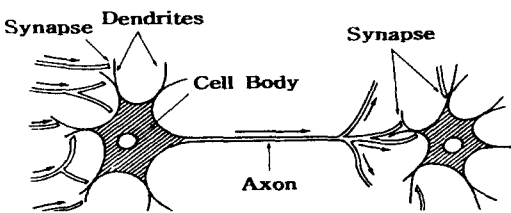


Fig. 1. Neuron.

이와 같이 인간의 두뇌는 신경세포가 매우 복잡하게 서로 연결된 비선형시스템의 네트워크라 할 수 있는데, 외부 또는 다른 신경세포로부터 시냅스(Synapse)를 통해 입력신호를 전달받아 세포체에서 적절한 처리를 하여 그 결과를 축삭돌기(Axon)를 통해 다른 신경세포에 전달하여 계산기능을 수행한다. 이 때 시냅스에서 두 세포

사이의 연결강도가 변화하며, 이 연결강도의 변화과정은 신경세포 사이의 영향력을 조절하는 뇌의 학습과정이며, 입력에 대한 신경회로망의 반응을 결정하는 역할을 하게 되는 것이다.

2.2. 인공신경망의 개요

인간두뇌에서 이루어지는 학습과정을 수학적으로 표현하여 모형화한 것을 자연신경망과 구별하여 인공신경망이라 부른다. 이 인공신경망을 구성하는 처리소자는 입력을 통합하는 결합함수와 통합된 결과의 동작범위를 제공하는 전달함수로 구성된다. 인공신경망은 수학적 연산능력을 가진 처리소자들이 상호연결된 것으로 적절한 학습규칙에 의해 동작한다.

이러한 신경망은 패턴분류 등 여러 가지 문제들을 처리함에 있어서 기존의 방법이 순차적이고 논리적인 틀속에서 프로그램되는 것과는 달리 학습을 통해 인지하게 되고, 본질적으로 다수의 유니트들에게 서로 영향을 주며 동시에 서로 다른 처리를 행하는 병렬처리를 한다는 특징이 있다.

뉴런을 정보처리소자라고 간주한다면 이는 다음 Fig. 2와 같이 n 개의 입력신호 x_1, x_2, \dots, x_n 을 받아 들이고 그것을 가지고 계산하여 출력 z 를 답으로 내 보내는 소자이다.

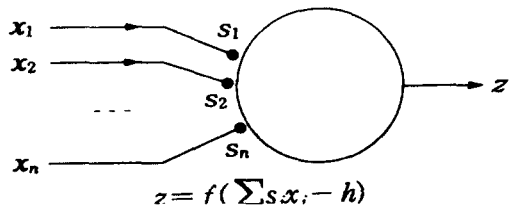


Fig. 2. Unit.

일반적으로 인공적인 뉴런의 모형을 Unit라 부르며, 다입력 1출력소자가 사용된다. 이러한 Unit는 다른 여러 개의 Unit와 연결되어 전체적인 인공신경망을 이루게 되며, 이들 Unit간은 연결강도에 의해 결합되어 있고 한 쪽 방향으로만 연결된다. 연결강도는 한 처리소자가 다른 처리소자에 미칠 수 있는 영향력의 크기를 나타낸다. 따라서 학습과정에서 주어진 입력에 대하여 출력과의 상관관계가 큰 영향력이 있는 입력층의 처리소자에 큰 연결강도를 부여하게 된다. 처리소자들은 여러 다른 처리소자들과 각각의 영향력에 따라 크기가 다른 연결강도로 연결되며 이로부터 전달되는 연결값을 변형시켜 다른 처리소자에 입력될 값을 만드는데, 이를 활성화값(Activation Value)이라 하며 전달함수에 의해 수행된다. 연결강도가 곱해진 입력치의 총합이 전달함수에 의해 변형을 받은 후 출력된다. 이 때 전달함수로는 Fig. 3과 같이 입력 x 에 대해 선형적으로 동작의 범위를 제공하여 출력하는 선형함수(Linear Function), x 가 X 를 넘을 때는 1을 출력하고 나머지는 0을 출력하는 계단함수(Step Function) 및 0과 1사이의 값을 출력하는 시그모이드 함수(Sigmoid

Function) 등이 주로 사용된다. 본 연구에서는 오차역전파 학습방법을 적용하기 위해서 미분가능한 형태의 비선형함수로 시그모이드 함수를 사용하였는데, 시그모이드 함수를 사용하면 뉴런의 출력이 아날로그 형태로 나타나는 장점이 있다.

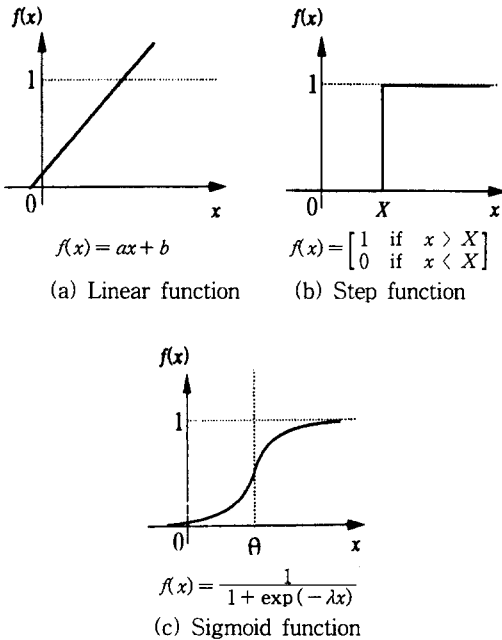


Fig. 3. Transfer functions.

그리고 출력치는 다음 식(1)을 이용하여 구한다.

$$z = f(\sum x_i s_i - h) \tag{1}$$

여기서, z 는 출력치, x_i 는 각각의 입력치, s_i 는 각각의 연결강도, h 는 임계치이다. 주어진 입력이 임계치 이상이 되면, 신경망은 활성화하여 출력을 내게 된다.

일반적인 인공신경망 모형은 다음과 같은 식(2)~(5)와 같이 간단히 나타낼 수 있다.

$$X = (x_1, x_2, \dots, x_n) : \text{입력 데이터} \tag{2}$$

$$S = (s_1, s_2, \dots, s_n) : \text{연결강도} \tag{3}$$

$$NET = x_1 s_1 + x_2 s_2 + \dots + x_n s_n \tag{4}$$

$$OUT = f(NET) : \text{결과(출력) 데이터} \tag{5}$$

2.3. 오차역전파 방법에 의한 학습

학습(Learning)은 인공신경망이 정보, 지식을 습득하는 방법으로 실제값과 출력값 사이의 오차를 줄이도록 반복실행하여 연결강도를 조정하는 것이다. 즉, 신경망에 있어서 연결강도는 임의의 초기값에서 시작하여 적절한 연결강도를 구하기 위하여 계속적인 개선이 이루어지게 되며 이 과정을 학습이라 한다.

현재 가장 많이 사용되고 있는 학습방법은 오차역전

파 방법인데, 크게 다음과 같은 두 단계로 구성되어 있다. 첫째 단계는 신경망의 입력패턴에 따라 각 처리소자의 입력치와 전달함수를 사용하여 출력을 산출하는 순방향처리이고, 두 번째 단계는 실제값과 계산치와의 차이인 오차를 역으로 진행시키면서 오차가 최소가 되도록 층과 층사이의 연결강도를 변화조정하는 역방향처리이다. 즉, 순방향처리는 자료의 입력전·후에 생성된 출력과 실제 목표값과의 오차를 계산하여 이를 중간층에 전파하는 과정이고, 역방향처리는 출력값과 실제값과의 오차를 최소화하도록 중간층과 출력층의 연결강도를 조절하여 중간층의 출력값을 구해 중간층과 입력층 사이의 연결강도를 반복적으로 조절하는 것이다. 이러한 학습은 오차를 최소화 하기 위한 과정으로 계속하여 반복실행된다.

오차역전파방법을 사용하여 신경망에 주어진 입력에 대해서 목적하는 출력을 하도록 하기 위해서, 먼저 신경망의 가중치에 대한 오차함수를 정의하고 이 오차함수를 최소화 하기 위한 방법으로써 도함수를 사용하는 경사법(Gradient Method)을 사용한다. 경사법은 예측치와 실측치간의 제곱오차로 표현되는 목적함수의 기울기를 구하여 매개변수의 사이에서 탐색이동과정을 반복함으로써 오차의 최고점에 도달하고, 이 때의 최적매개변수를 구하는 방법이다. 이 때 오차역전파방법의 오차함수는 다음 식과 같다. 패턴 p 가 입력되었을 때의 출력패턴의 k 번째 성분을 O_{pk} , 목적패턴의 k 번째 성분을 t_{pk} , 오차함수를 $E_p(s)$ 라 할 때

$$E_p(s) = \frac{1}{2} \sum_k (t_{pk} - O_{pk})^2 \tag{6}$$

가 되고, p 개의 패턴에 대한 평균오차함수를 $E(s)$ 라 하면

$$E(s) = \sum_p E_p = \frac{1}{2p} \sum_p \sum_k (t_{pk} - O_{pk})^2 \tag{7}$$

가 된다. 이 오차함수가 최소값을 가지는 모든 연결가중치의 값을 구하는 것이 학습과정이다.

다음 Fig. 4는 본 인공신경망 모형의 전체적인 수행절차이다.

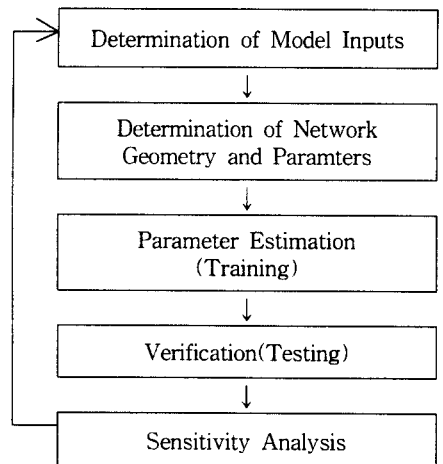


Fig. 4. Logic diagram of Artificial Neural Network Model.

3. 모형의 적용 및 분석

3.1. 대상유역 및 분석자료

본 연구에서 인공신경망기법을 이용한 모형을 적용하기 위한 대상하천으로는 한반도의 동남단에 위치한 낙동강을 선택하였다(Fig. 5). 낙동강은 강원도 태백산에서 발원하여 경상북도와 경상남도를 관류하고 부산광역시 사하구 하단지점에서 남해로 유입하는 우리나라 제2의 하천이다. 낙동강 수계는 총연장 521.5km에 825개의 소하천으로 구성되어 있어 다양한 오염원에 의한 수질오염 사고가 발생할 가능성이 매우 높은 유역이다.

따라서 낙동강 본류에서 인공신경망기법에 의한 모형을 적용하기 위한 대상지점으로는 낙동강 상류수질을 보여주는 상주1 지점(상풍교)과 낙동강 중류수질을 보여주는 달성지점(강정취수장)을 선택하였다. 특히 달성지점은 대구광역시의 주요한 취수원인 강정취수장이 위치하고 있는 지점이어서 의미가 있다고 생각된다.

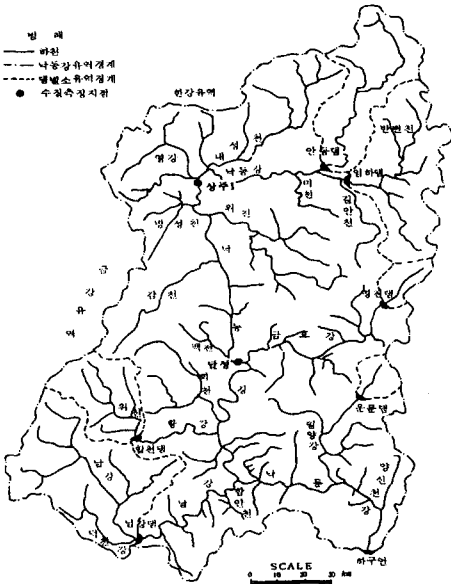
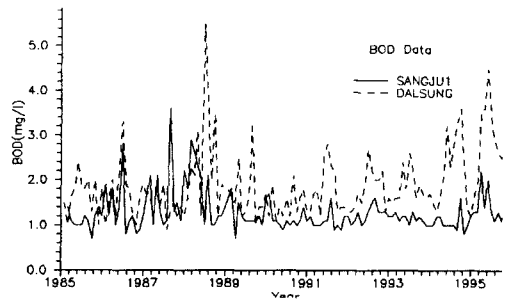


Fig. 5. Study river.

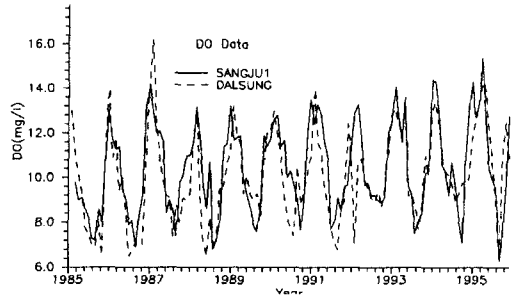
본 연구에 필요한 자료는 강우량, 유량 및 수질자료이며, 대상 지점별로 1985년에서 1995년까지의 자료를 이용하였다.^{18),19)} 여기서, 강우량 자료는 상주1 지점에 대해서는 안동기상대, 달성지점에 대해서는 대구기상대 자료를 이용하였고, 유량자료는 상주1 지점에 대해서는 사벌관측소, 달성지점에 대해서는 성주관측소 자료를 이용하였으며, 수질자료는 환경통계연감의 상주1 및 달성지점의 자료를 이용하였다(Fig. 6).

여기서, 인공신경망기법에 의한 모형수행을 위한 학습 자료는 1985년부터 1993년까지의 9개년에 걸친 월별 강우량, 유량 및 수질자료를 이용하였으며, 학습에 사용되지 않은 1994년부터 1995년까지의 월별 강우량, 유량 및 수질자료(BOD, DO)는 본 모형에 대한 적용성을 보기 위한 예측자료로 사용하였다. 반복횟수는 500회에서

5,000회 정도까지 변화시켜 가면서 적절한 값을 구하고자 하였다.



(a) BOD data



(b) DO data

Fig. 6. Study data.

3.2. 인공신경망 모형의 구성

수질예측을 위한 인공신경망 모형은 다음 Fig. 6과 같이 입력층(Input Layer), 은닉층(Hidden Layer) 및 출력층(Output Layer)으로 구성된 전형적인 다층신경망 구조로 작성하였다. 여기서 입력층은 다른 층으로부터 들어오는 연결선 없이 신경망 외부로부터 입력을 받아 전달하게 되며 외부의 입력값을 그대로 은닉층으로 전달시킨다. 따라서 은닉층은 입력층과 출력층 사이에 연결되어 각 패턴마다 유사한 특성에 따라 필요한 정보를 학습시켜 저장시킬 수 있도록 한 것이다. 출력층은 은닉층에서 출력된 값을 받아 신경망 전체의 최종 출력으로 사용한다. 각 층은 일종의 여과기능(Filtering)을 해주는 것으로

Input Layer Hidden Layer Output Layer

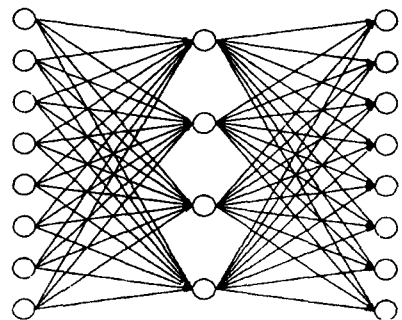


Fig. 7. Structures of neural networks.

신경망의 입력신호를 받아 처리하여 출력패턴을 발생시키며, 불완전한 입력에 대해서도 매개변수의 인식에 의해 출력값을 생성시킬 수 있다. 그러므로 불완전한 입력 신호가 들어와도 이전에 학습된 입력신호와 유사하다고 판단되면 신경망이 인식하고 있는 출력패턴을 발생시킨다. 또한 다층신경망의 각 층에 있는 모든 처리소자들은 대부분 방향성을 가진 연결선으로 연결되어 있고 각 연결선들은 연결강도를 가지고 있어 주어진 입력에 따라 그 연결선을 거쳐 출력된다.

학습자료는 입력층을 강우량, 유량 및 수질자료에 대해 선행 2개월에 걸친 자료로 구성하는 방법 I 과 선행 4개월에 걸친 자료로 구성하는 방법 II, 선행 6개월에 걸친 자료로 구성하는 방법 III의 세 가지로 하였다. 이에 따른 수질예측을 위한 신경망 모형의 입력층, 은닉층 및 출력층의 구조는 다음 Table 1과 같다. 이 때 은닉층에 대하여 Hertz 등(1991)은 입력층이 n 이라면 $2n$ 또는 $2n+1$ 로 충분하다고 하였는데, 본 논문에서는 이중 $2n+1$ 을 선택하였다.

Table 1. Numbers of layer for each method

Method	Layer		
	Input	Hidden	Output
I	6	13	1
II	12	25	1
III	18	37	1

이 때 신경망의 출력치는 변환함수인 시그모이드 함수를 통해 0에서 1 사이의 값으로만 주어지기 때문에 적절한 정규화(Normalization) 과정이 필요하다. 일반적인 정규화 과정은 다음과 같이 식(8)을 이용한다.

$$T_{ar} = T_{min} + \left[\frac{V_{ar} - V_{min}}{V_{max} - V_{min}} (T_{max} - T_{min}) \right] \quad (8)$$

여기서, V_{max} 와 V_{min} 은 변환되기 전 출력값의 최대값과 최소값이고, T_{max} 와 T_{min} 은 변환된 출력값의 최대값과 최소값으로 일반적으로 0.9와 0.1을 사용하며, T_{ar} 은 변환된 출력값, V_{ar} 은 변화되기 전의 출력값으로 T_{ar} 함수의 입력값이다.

여기서, 오차역전과 방법에 의한 학습과정을 단계별로 설명하면 다음과 같다.

- ① 1단계 : 인공신경망의 상태를 결정하는 연결강도 S_{ij} , S_{kj} 와 초기값 θ_j , θ_k 를 각각 작은 값의 임의의 수로 초기화 한다.
- ② 2단계 : 학습패턴을 방법별로 설정한다.
- ③ 3단계 : 은닉층 유니트 j 의 출력을 계산한다. 입력층과 은닉층 사이의 연결강도 S_{ij} 와 은닉층 유니트 j 의 초기값 θ_j 를 이용해 j 의 입력 NET_{ij} 를 구해 시그모이드 함수 f 로 j 의 출력 O_{ij} 를 구한다.
- ④ 4단계 : 출력층 유니트 k 의 출력을 계산한다. 은닉층과 출력층 사이의 연결강도 S_{kj} 와 출력층 유니트 k 의 초기값 θ_k 를 이용해 k 의 입력 NET_{pk} 를 구해 시그모이드 함수 f 로 k 의 출력 O_{pk} 를 구한다.
- ⑤ 5단계 : 출력층 유니트의 오차를 계산한다. 학습패

턴의 목표출력 t_{pk} 와 실제출력 O_{pk} 와의 차로부터 출력층 유니트 k 에 연결된 연결강도와 k 의 초기값에 대한 오차 δ_{pk} 를 구한다.

- ⑥ 6단계 : 은닉층 유니트의 오차를 계산한다. 오차 δ_{pk} 와 은닉층과 출력층간의 연결강도 S_{kj} 와 은닉층 출력 NET_{ij} 로부터 j 에 연결된 연결강도와 은닉층 유니트의 오차 δ_{ij} 를 구한다.
- ⑦ 7단계 : 은닉층과 출력층 사이의 연결강도와 출력층 유니트의 초기값을 변경한다.
- ⑧ 8단계 : 입력층과 은닉층 사이의 연결강도와 은닉층 유니트의 초기값을 변경한다.
- ⑨ 9단계 : 다음 패턴을 학습한다.
- ⑩ 10단계 : 학습패턴을 모두 학습할 때까지 2단계로 되돌아 간다.
- ⑪ 11단계 : 학습 반복횟수를 계산한다.
- ⑫ 12단계 : 학습 반복 종료횟수보다 작으면 2단계로 되돌아 간다.
- ⑬ 13단계 : 종료한다.

3.3. 적용결과의 비교분석

1) BOD

전절에서 구성된 인공신경망 모형을 적용하기 위하여 상주1 및 달성지점을 대상으로 1985년부터 1993년까지의 월별 강우량, 유량 및 BOD 자료를 이용하여 각 방법별로 학습을 수행하였으며, 수질자료에 대해 계산치와 관측치를 비교하였다. Fig. 8은 달성지점에 대해 방법 III으로 학습한 결과이다(BOD). 분석 결과 계산치와 관측치가 매우 근접한 결과를 나타내어 수질예측의 상관성이 높음을 알 수 있었다.

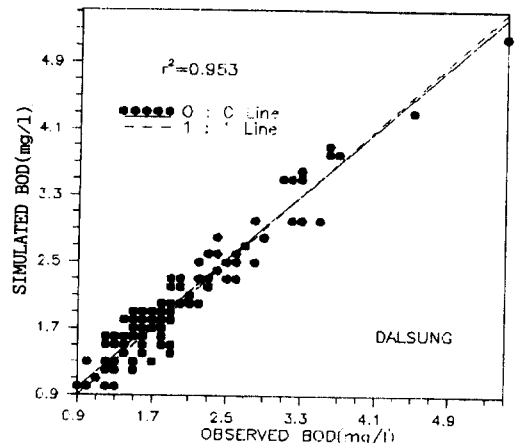


Fig. 8. Comparisons of observed and simulated BOD (Dalsung, Method III).

또한 위에서 학습된 인공신경망 모형을 1994년부터 1995년까지의 월별유량 및 BOD 자료에 대해 각 방법별로 적용하였으며, 그 결과는 Fig. 9와 같다. 분석 결과 상주1 지점에 대해서 방법 II를 수행하였을 때 1995년 4

월 등에서 추세와는 상이한 이상 결과가 나왔으나, 대체적으로 각 지점에서 실측치와 계산치 사이에 일치하는 예측결과를 나타내었고 상관성이 높은 것으로 파악되었다(Table 2 참조).

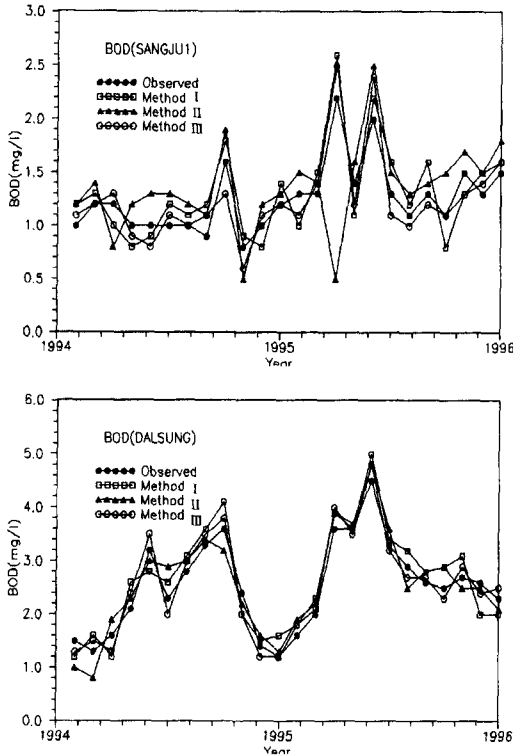


Fig. 9. Results of forecasting of BOD by each method.

2) DO

전절에서 구성된 인공신경망 모형을 적용하기 위하여 상주1 및 달성지점을 대상으로 1985년부터 1993년까지의 월별 강우량, 유량 및 DO 자료를 이용하여 학습을 수행하였으며, DO 자료에 대해 계산치와 관측치를 비교하였다. 그 결과 Fig. 8과 비슷한 유형을 나타내어 계산치와 관측치가 매우 근접한 결과를 나타내었으며, 수질예측의 상관성이 높음을 알 수 있었다.

또한 위에서 학습된 인공신경망 모형을 1994년부터 1995년까지의 월별 강우량, 유량 및 DO 자료에 대해 적용하였으며, 그 결과는 Fig. 10과 같다. 분석 결과 달성지점에 대해서 방법 I을 수행하였을 때 1994년 5월, 방법 II를 수행하였을 때 1995년 3월 등에서 추세와는 상이한 이상 결과가 나왔으나, 대체적으로 각 지점에서 실측치와 계산치 사이에 일치하는 예측결과를 나타내었고 상관성이 높은 것으로 파악되었다.

3) 적합성 검증

예측결과와 적합성을 판정하기 위하여 통계적인 검정을 실시하였다. 전절의 계산치와 관측치를 비교하기 위

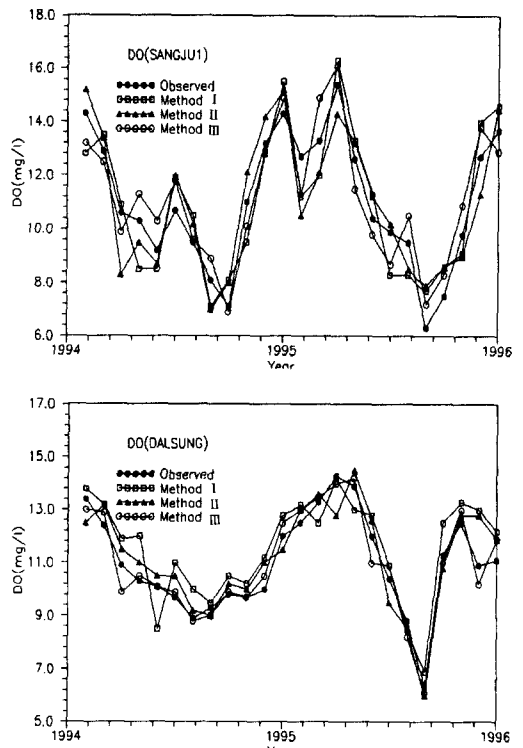


Fig. 10. Results of forecasting of DO by each method.

하여 모형의 효율도(Model Efficient, ME), 절대오차의 평균(Mean Square Error, MSE), 상대오차의 평균을 나타내는 Bias 및 체적오차의 백분율(Volume Error, VER)을 검토하였다. 그리고 이를 적합성 검증에 사용되는 방법들을 식으로 표시하면 아래 식(9)~(12)와 같다.

$$ME = 1 - \frac{\sum(Q_{obs} - Q_{est})^2}{\sum(Q_{obs} - \bar{Q}_{est})^2} \quad (9)$$

$$MSE = \left[\frac{\sum(Q_{obs} - Q_{est})^2}{n} \right]^{1/2} \quad (10)$$

$$Bias = \frac{\sum(Q_{obs} - Q_{est})}{n} \quad (11)$$

$$VER = \frac{\sum Q_{obs} - \sum Q_{est}}{\sum Q_{obs}} \times 100 \quad (12)$$

여기서, n은 자료의 수, Q_{obs}는 관측치, Q_{est}는 계산치이다.

Table 2(a). Results of the statistical test(Sangju 1)

Factor	Method	ME	MSE (mg/l)	Bias (mg/l)	VER (%)
BOD	I	0.890	0.224	-0.058	-4.636
	II	0.929	0.429	-0.100	-7.947
	III	0.942	0.173	0.008	0.662
5DO	I	0.899	1.135	0.029	0.264
	II	0.928	1.137	-0.004	-0.038
	III	0.942	0.930	-0.133	-1.208

Table 2(b). Results of the statistical test(Dalsung)

Factor	Method	ME	MSE (mg/l)	Bias (mg/l)	VER (%)
BOD	I	0.891	0.352	-0.125	-4.926
	II	0.922	0.302	-0.046	-1.806
	III	0.945	0.232	-0.029	-1.149
DO	I	0.892	1.000	-0.575	-5.231
	II	0.910	0.755	-0.213	-1.933
	III	0.949	0.487	0.004	0.038

그리고 결과의 적합성을 판정하기 위하여 검정을 실시한 결과 Table 2와 같이 나타났다. 먼저, 상주1 지점에서 BOD는 각 방법에 대하여 ME는 0.890~0.942, MSE는 0.173~0.429 mg/l, Bias는 -0.100~0.008 mg/l, VER는 -7.947~0.662%를 나타내었다. DO는 각 방법에 대하여 ME는 0.899~0.942, MSE는 0.930~1.137 mg/l, Bias는 -0.133~0.029 mg/l, VER는 -1.208~0.264%를 나타내었다. 달성지점에서 BOD는 각 방법에 대하여 ME는 0.891~0.945, MSE는 0.232~0.352 mg/l, Bias는 -0.125~-0.029 mg/l, VER는 -4.926~-1.149%를 나타내었다. DO는 각 방법에 대하여 ME는 0.892~0.949, MSE는 0.487~1.000 mg/l, Bias는 -0.575~0.004 mg/l, VER는 -5.231~0.038%를 나타내었다.

이와 같이 Fig. 7~8 및 Table 2의 결과를 비교해 볼 때 본 인공신경망 모형은 추세와는 다른 몇몇 이상 결과에도 불구하고, 대체적으로 계산치와 관측치가 매우 근접한 결과를 나타내었으며 두 결과의 상관성이 높음을 알 수 있었다. 또한 3가지 방법을 비교해 보면 방법 III이 방법 I 및 방법 II에 비해 실측치에 더욱 근사한 것으로 나타났다. 그 이유는 방법 III이 다른 방법들에 비해 더욱 구체적인 자료(선행 6개월의 강우량, 유량 및 수질자료)로 구성되어 학습에서 다른 방법에 비해 추세를 더 정확히 반영하였기 때문이라고 판단된다.

4. 결론

본 연구에서는 낙동강유역의 하천수질 관리를 위한 하천유지용수 공급계획, 하수처리장 신설 및 증설계획과 같은 중·장기적인 수질관리 대책을 수립함에 있어서 사용될 수 있는 인공신경망기법을 이용한 수질예측 모형을 개발하고, 상주1 및 달성지점에 대해 인공신경망 모형을 적용하여 계산치와 관측치를 비교 검토하여 그 적용성을 확인하였다.

1. 낙동강 수계의 상주1 및 달성지점을 대상으로 수질예측을 수행할 수 있는 인공신경망 모형을 입력층, 은닉층 및 출력층으로 구성하고 1985년부터 1993년까지의 자료를 가지고 오차역전과 방법에 의해 학습을 수행하였다.

2. 학습한 인공신경망기법을 이용한 수질예측 모형을 1994년부터 1995년까지의 월별유량 및 수질자료에 대해 적용한 결과 계산치와 실측치의 도시적인 비교와 통계적인 검정을 통해 볼 때, 본 모형은 수질실측치에 대해 매

우 상관성이 있는 결과를 얻은 것으로 나타났다.

3. 인공신경망 모형에서 여러 가지 경우(방법 I, II 및 III)를 제시하였으며, 자료가 구체적(입력층의 처리소자의 수가 많은 것)이어서 관측치의 추세를 좀더 잘 반영한 방법 III이 실측치에 더욱 가까운 결과를 얻었다.

이상과 같이 하천수질 관리를 위한 수질예측을 위하여 제시한 인공신경망 모형을 적용 분석한 결과 그 효용성을 확인할 수 있었으며, 학습자료를 더 축적한다면 더욱 신뢰성 높은 수질예측 결과를 얻을 수 있을 것으로 판단된다.

참고 문헌

- Huck, D. M., G. J. Farquhar, 1974, Water quality models using the Box-Jenkins method, J. Envir. Eng. Div., ASCE, EE5, 110(3).
- Mehta, B. M., R. C.Ahlert and S. L. Yu, 1975, Stochastic variation of water quality of the Passaic River, Water Resources Research, 11(2), 300~308.
- Thomann, R. V., 1976, Time series analysis of water quality data, J. Sanit. Eng. Div., ASCE, 93(1), 1~23.
- Chadderton, R. A., A. C. Miller and McDonnel, A. J., 1982, Uncertainty analysis of dissolved oxygen model, J. Envir. Eng. Div., ASCE, EE5, 1003~1011.
- Donigian, A. S., Jr. and W. C. Huber, 1991, Modeling nonpoint source water quality in urban and non-urban areas, EPA/600/3-91-039, U.S. Environ. Protection Agency, Athens, GA.
- Wanielista M. P., Y. A. Yousef and W. M. McLellon, 1997, Nonpoint source effects on water quality, JWPCF, 49, 441~451.
- 권혁윤, 1992, 하천 유량변동에 따른 수질변화에 관한 연구, 연세대학교 보건대학원 석사학위논문.
- 유희정, 1993, 수문인자에 의한 하천수질변화에 관한 연구, 한국수문학회지, 26(3), 75~85.
- 신현석, 윤용남, 1993, SWMM 모형을 이용한 도시유역의 유출 및 NPS 오염물질 배출모의, 한국수문학회지, 26(3), 125~135.
- 이상수, 1994, ARIMA 모형을 이용한 금강의 장래 수질예측에 관한 연구, 중앙대학교 대학원.
- 양우석, 1994, 하천유량 변화에 따른 수질오염도 분석, 수원대학교 산업경영대학원 석사학위논문.
- 백경원, 김상호, 한건연, 송재우, 1995, QUAL2E에 의한 한강하류부에서의 수질해석, 대한토목학회 논문집, 15(2).
- 이상도, 1996, 중소하천의 유량과 수질변동 특성에 관한 연구, 홍익대학교 대학원 석사 학위 논문.
- 김준수, 1998, 남강댐 유역의 하천수질 예측에 관한 연구, 경상대학교 대학원 석사학위 논문.
- Holger R. Maier and Graeme C. Dandy, 1996, The use of artificial neural networks for the prediction of water quality parameters, Water Resources

- Research, 32(4), 1013~1022.
- 16) 강두선, 신현석, 김중훈, 1998, 신경망 모형을 이용한 장기하천수질의 예측, 1998년도 대한토목학회 학술 발표회, 23~26.
- 17) 신현석, 최시중, 김중훈, 1998, 신경망을 이용한 도시 유역 유출 및 비점원 오염원 배출 모형화 연구, 대한 토목학회 논문집, 18(5), 437~448.
- 18) 환경통계연감, 환경부.
- 19) 수문조사연보, 건설교통부.