

## 실시간 수질 예측을 위한 신경망 모형의 적용

조용진 · 연인성<sup>\*\*</sup> · 이재관<sup>\*\*</sup>

충주대학교 환경공학과

<sup>\*</sup>충북대학교 토목공학과

<sup>\*\*</sup>국립환경연구원

## Application of Neural Network Model to the Real-time Forecasting of Water Quality

Yong-Jin Cho · In-Sung Yeon<sup>\*\*</sup> · Jae-Kwan Lee<sup>\*\*</sup>

*Dept. of Environmental Engineering, Chungju National University,*

*<sup>\*</sup>Dept. of Civil Engineering, Chungbuk National University*

*<sup>\*\*</sup>National Institute of Environmental Research*

*(Received 9 January 2004, Accepted 31 May 2004)*

**Abstract** : The objective of this study is to test the applicability of neural network models to forecast water quality at Naesa and Pyongchang river. Water quality data divided into rainy day and non-rainy day to find characteristics of them. The mean and maximum data of rainy day show higher than those of non-rainy day. And discharge correlate with TOC at Pyongchang river. Neural network model is trained to the correlation of discharge with water quality. As a result, it is convinced that the proposed neural network model can apply to the analysis of real time water quality monitoring.

**keywords** : Neural network, Discharge, Water quality, Real-time water quality monitoring

### 1. 서 론

수질 오염을 발생시키는 오염원으로는 인구, 가축, 산업 폐수, 토지이용 상태 및 인위적인 비료사용 등을 들 수 있다. 하수처리율이 향상되고 경제활동 수준이 증가함에 따라 강우시 유출되는 비점오염원 부하가 수질에 미치는 영향은 높아지고 있다. 강우시에는 수질오염 사고가 발생할 가능성이 많으며 특히 충주호는 상수원으로 활용되고 있어서 오염 사고에 대한 대비가 요구된다(조, 2002).

1994년 낙동강 수질오염 사고를 계기로 환경부에서는 4대강 유역에 수질자동측정망을 설치하여 운영하고 있다. 주요 상수원 및 오염사고 취약지역 등에 설치되어 수질상태를 상시 감시 기능을 담당하고 있으나 효과적인 운영측면에서 측정망수의 지속적인 확충이 필요하다. 현재 한강유역에는 양평, 가평, 경안천, 구리, 원주, 여주, 평창강, 달천의 8개소가 관리, 운영되고 있다. 충주호 상류에 위치한 평창강 수질자동 측정지점은 가장 최근에 설치되었으며 수질측정 항목은 수온, pH, DO, TOC, 물고기 독성, TN, TP 등이다(김, 2003).

현재 수질자동측정망 지점에서 수질오염 사고에 대한 판단으로 물벼룩의 상태 등을 이용하고 있으나 수질관리 차원에서 보다 효과적인 운영을 위해서는 적합한 예측 모형과 연계하여 관리자의 판단을 도울수 있는 방안이 요구된다.

예측 모형은 실시간으로 수집되는 자료를 효율적으로 활용할 수 있어야 하고 입력 자료를 최소화하여 적절한 시간 범위 내에서 예측을 위한 연산이 이루어질 수 있어야 한다. 또한 여건의 변화에 스스로 적응해 갈 수 있는 모형이어야 한다. 이러한 측면에서 볼때 적절한 모형 중에 신경망 모형을 들 수 있다. 신경망 모형은 입력과 출력만을 고려하여 구성이 용이하며 입력자료의 변화에 따른 학습을 통해 예측값을 출력하기 때문에 변화에 잘 적응하여 실시간 자료의 적용에 적합한 모형이라 판단된다.

또한 강우로 인해 유량의 변화가 큰 강우시와 유량변동이 작은 무강우시는 유량변화 뿐만아니라 수질변화 특성이 다르다. 오염원을 이송하는 것은 유량입을 감안할 때 강우의 유무에 따른 유량과 수질 관계의 해석은 매우 중요하다(이, 1995). 유량 측정시간과 수질 측정시간이 다르거나 동일 시간에 측정된 자료의 부족으로 실시간 수질변화에 대한 명확한 관계 규명에 어려운 점이 있었다.

이 연구에서는 내사천 유역과 평창강 유역을 중심으로 강우시와 무강우시의 수질특성을 분석하였으며 강우시에 측정된 수질과 유량 자료를 이용하여 실시간 수질 예측을 위한 신경망 모형의 적용성을 검토하고자 한다.

### 2. 연구내용 및 방법

#### 2.1. 대상유역

대상유역은 충주호 유역 내에 있는 내사천 유역과 평창

<sup>\*</sup> To whom correspondence should be addressed.  
hifivee@netian.com

강 유역이다. 내사천 유역의 상류는 전형적인 산림지역이고, 중류에 마을이 형성되어 있으며, 최하류는 충주호로 합류된다.

내사천 유역의 면적은 2.4 km<sup>2</sup>이고 내사천 하류의 하천폭은 5~6m인 소하천 지역이다. Fig. 1에 내사천 유역을 나타내었다.

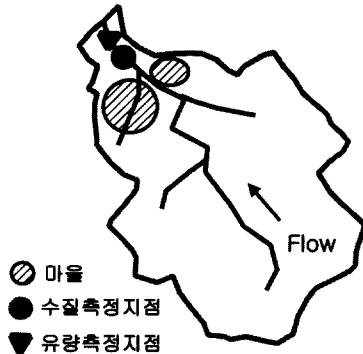


Fig. 1. Naesa basin.

평창강 유역은 유로연장 145 km, 유역면적 1781 km<sup>2</sup>에 해당하는 지방1급 하천이다. 평창강은 강원도 영월군 서면에서 주천강과 합류한 후 강원도 영월군 영월읍에서 남한강 본류와 합류한다.

평창강 유역내에 위치한 하류부의 유량측정 지점은 영월 1 지점으로서 남한강 본류와 합류하는 최하류 지점으로부터 상류 약 14 km 지점에 위치하고 있다. 수질자동 측정망이 운영되고 있는 곳은 충북 제천시 송학면 장곡리로서 평창강과 주천강이 합류하는 지점의 직하류에 위치하고 있으며 유량측정 지점으로부터 약 11 km의 상류 지점이다.

유량측정 지점과 수질자동 측정지점 사이에는 지천이 유입되고 있다. 그러나 평창강 전체 유역면적에 비하여 지천의 유역면적이 매우 작아서 평창강 본류의 유량변화에 미치는 영향은 미미한 것으로 판단된다. Fig. 2에 평창강 유역을 나타내었다.



Fig. 2. Pyongchang river basin.

## 2.2. 수질조사 및 분석

내사천 유역에서 강우시와 무강우시로 구분하여 수질을 조사하였으며 강우시에는 매시간마다 시료 채취와 유량측정이 병행되었다. 무강우시에는 하천유량과 수질이 비교적 큰 차이없이 일정함에 따라 1회 출수시 4시간~6시간 간격으로 최소 24시간 동안 시료 채취와 유량측정을 실시하였다.

수질조사 항목은 BOD, TN, TP를 중심으로 분석하였으며 분석은 수질오염공정시험법과 Standard Method를 기준으로 하였다.

평창강 유역에서의 수질자료는 환경부에서 설치한 평창강 지점의 수질자동 측정망 자료를 이용하였다. 연구에서는 5분마다 측정된 자료로부터 시간위로 환산한 TOC 자료를 이용하였다.

유량자료는 T/M 수위국인 영월1 지점의 자료를 이용하였다. 시간위로 계측된 수위를 수위-유량 관계곡선식(한국수자원공사, 2002)을 이용하여 시간위 유량으로 환산하였다. 유량변화에 따른 수질변화 및 상관성을 파악하기 위해 표준편차(Std-Dve.), 상관계수(CC)를 이용하였으며 모형의 적용결과를 분석하기 위해 제공된 평균오차(RMSE) 등을 이용하였다.

## 2.3. 수질예측을 위한 신경망 모형

신경망 모형에서 가장 많이 이용되는 다층신경망 모형은 주로 오차 역전파 알고리즘을 이용하여 학습을 수행하는데 입력패턴을 이용해 네트워크가 계산한 출력값과 목표값과의 오차를 최소화시키는 방향으로 연결강도를 조절하는 방법이다. 네트워크는 입력층, 은닉층, 출력층 방향으로 연결되어 있으며 출력값을 생성하는 전달함수는 시그모이드 함수가 많이 사용된다.  $u_y$ 를 출력층의 가중합이라 하고, 경사도  $\lambda=1$ 인 단극성 시그모이드 함수를 전달 함수( $\phi$ )로 사용하는 경우 출력  $Y$ 는 식 (1)과 같다.

$$Y = \phi(u_y) = \frac{1}{1 + \exp(-u_y)} \tag{1}$$

학습하는 동안 목표값과 실제 출력값 사이의 계산된 오차 정보는 네트워크를 통하여 역방향으로 전달되며 연결가중치에 영향을 준다.  $p$  학습단계에서의 은닉층과 출력층간의 연결강도 변화량을  $\Delta W^p$ , 입력층과 은닉층간의 연결강도 변화량을  $\Delta V^p$ 라 할 때  $p+1$  학습단계의 조정된 연결강도  $W^{p+1}$ 와  $V^{p+1}$ 는 식 (2), 식 (3)과 같다.

$$W^{p+1} = W^p + \Delta W^p \tag{2}$$

$$V^{p+1} = V^p + \Delta V^p \tag{3}$$

계속적으로 연결강도를 변경하며 학습회수에 만족을 하거나 오차가 허용범위에 수렴하면 학습을 종료하게 되며

학습을 마친 연결강도를 사용하여 예측을 수행하게 된다 (안 등, 2000). 수질예측을 위한 신경망 모형은 연결강도 조정시에 BP 알고리즘에 비해 우수성이 입증된 Levenberg-Marquardt 방법을 사용하였다.

Table 1에 적용한 신경망 모형의 입출력 구성을 나타내었다. 강우로부터 발생된 t시간의 유량자료와 t시간의 수질 자료, 강우발생후 경과시간을 입력자료로 하여 t+1시간에서의 수질을 출력하도록 하였다. 정량적 자료인 유량과 수질 자료만을 학습하여 출력하는 LM(Levenberg-Marquardt) 신경망 모형과 정성적 자료인 강우발생 경과시간을 추가하여 학습한 MD(Modula) 신경망 모형을 적용하여 비교하였다.

Table 1. Construct of Neural network

Model	Input 1	Input 2	Input 3	Output
LM	Water quality (t)	Discharge (t)	-	Water quality (t+1)
MD	Water quality (t)	Discharge (t)	Time of rainfall process	Water quality (t+1)

입력자료에 강우발생경과시간을 나타내는 정성적인 자료를 첨부한 MD 신경망 모형은 일반적인 신경망 모형과 달리 은닉층을 분리하여 정량적인 자료와 정성적인 자료를 구분하여 학습할 수 있는 장점이 있다. LM 신경망 모형과 MD 신경망 모형의 구조를 Fig. 3에 나타내었다.

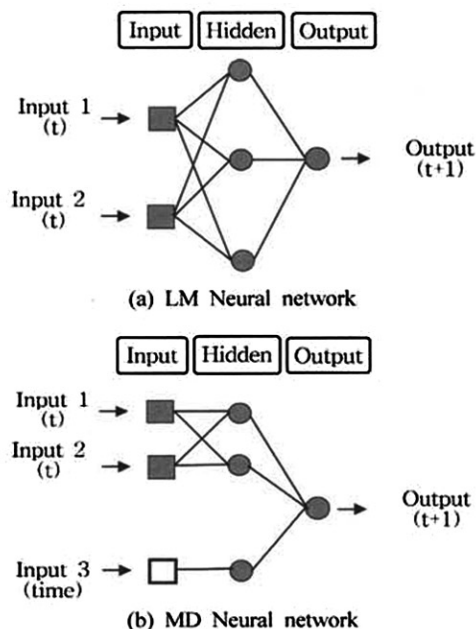


Fig. 3. Structure of Neural network.

### 3. 결과 및 고찰

#### 3.1. 조사자료

##### 3.1.1. 내사천 유역

강우시의 자료는 2002년 7월 5일과 10월 5일에 측정되었

으며 무강우시 자료는 2002년 5월 24일과 6월 19일, 10월 20일에 측정되었다. 소하천인 내사천 유역에서 강우시와 무강우시에 측정된 수질 및 유량(Q)의 특성을 Table 2에 나타내었다.

강우시의 BOD, TN, TP 수질 측정값은 모두 무강우시에 비해 평균, 최대값이 크게 나타났다. 특히 TP의 최대값은 강우시 4.010 mg/L 무강우시 0.172 mg/L로 매우 큰 차이를 보이고 있으며 마을 주변 경작지에서의 비료 사용이 큰 원인으로 판단된다.

또한 강우시 측정된 수질자료들의 표준편차는 유량변화가 작은 무강우시에 비해 매우 크게 나타나 유량의 증가에 따른 수질의 변화는 큰 것으로 분석된다.

Fig. 4에 내사천 하류 지점에서 강우시 측정된 유량을 나타내었으며 Fig. 5에 측정된 BOD, TN, TP의 농도를 Log 축으로 하여 도시하였다.

Table 2. Statistical analysis of water quality at Naesa

Index	Statistics	BOD (mg/L)	TN (mg/L)	TP (mg/L)	Q (cmd)
Rainy day	Min	0.60	1.45	0.090	1062.12
	Mean	1.94	5.62	1.080	15647.04
	Max	5.80	16.20	4.010	76833.88
	Std-Dev.	1.05	3.63	0.872	15165.81
Non-rainy day	Min	1.40	2.70	0.090	852.12
	Mean	1.70	4.25	0.131	956.23
	Max	2.00	5.13	0.172	1053.15
	Std-Dev.	0.31	1.35	0.041	100.68

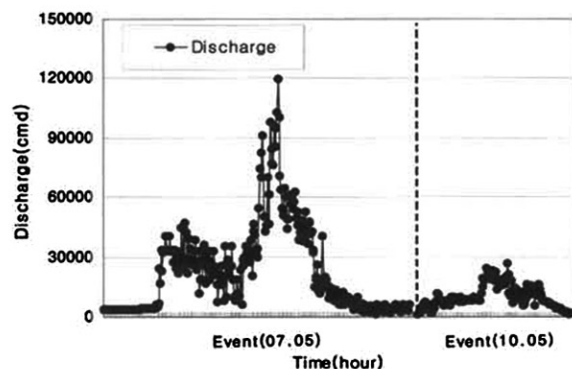


Fig. 4. Discharge of at Naesa.

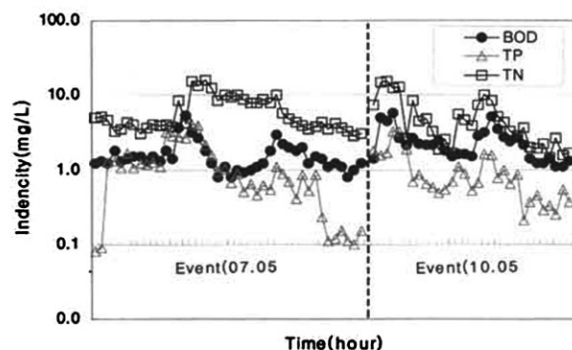


Fig. 5. Water quality of at Naesa.

2002년 7월 5일 사상의 경우 최대 강우가 발생했던 시점은 7월 5일 20시였으며 침투 유량이 발생한 시점은 7월 5일 21~22시로 나타났다. 10월 5일 사상도 유사한 결과를 보였으며 유역면적이 작고 하상 경사가 심한 내사천에서도 도달시간은 1~2시간으로 나타났다.

강우발생에 따라 유량이 급격히 증가하는 시점에서 수질이 큰 폭으로 변화하였으며 유량 증감에 따라 수질이 변동하는 경향을 보였다. 강우 초기에 노출되어 있는 비점오염원의 유입으로 수질이 크게 변동하였으며 유입 유량에 의해 큰 영향을 받는 것으로 판단된다.

3.1.2. 평창강 유역

Table 3에 평창강 유역에서 취득된 수온(Temp.), TOC, 수위(H), 유량(Q)의 통계특성을 나타내었다. 2003년 1월, 4월, 7월 시단위 자료이며 결측된 부분을 제외하였다.

Table 3. Statistical analysis of TOC at Pyongchang river

Index	Statistics	Temp. (°C)	TOC (mg/L)	H (m)	Q (cms)
Rainy day	Min	10.20	0.60	0.58	22.72
	Mean	17.50	1.29	2.11	304.29
	Max	23.40	2.20	4.75	1308.94
	Std-Dev.	3.16	0.41	0.87	261.39
Non-rainy day	Min	0.40	0.30	0.20	3.73
	Mean	6.82	0.62	0.44	16.29
	Max	24.10	0.90	0.98	55.35
	Std-Dev.	8.34	0.11	0.23	13.97

강우시 TOC는 무강우시에 비해 평균, 최대값이 크게 나타났으며 TOC의 최대값은 강우시 2.20 mg/L, 무강우시 0.90 mg/L로 큰 차이를 보이고 있다.

강우시에는 표준편차 또한 커서 수질의 변동이 큰 것으로 나타났으며 내사천의 결과와 유사한 것으로 분석되었다.

강우시와 무강우시의 자료를 1월, 4월, 7월의 월별로 구분하여 분석하였으며 월별 TOC의 평균값을 Table 4에 나타내었다.

Table 4. Statistical analysis of TOC at Pyongchang river (mg/L)

Month	Statistics	Rainy	Non-rainy	Total
1	Mean	-	0.60	0.60
4	Mean	1.20	0.60	0.70
7	Mean	1.40	0.70	0.80
Total	Mean	1.30	0.60	0.70
	Std-Dev.	0.41	0.11	0.28

평창강 유역의 TOC 평균값에 대한 월별 특성을 보면 여름철인 7월에 0.8 mg/L로 가장 크고 다음으로 봄철인 4월에 0.7 mg/L, 겨울철인 1월에 0.6 mg/L로 나타났다. 그러나 2003년 1월의 강우는 미미하여 유량에 영향을 미치지 않는 것으로 판단된다. 4월, 7월 모두 강우시 TOC의 평균값이

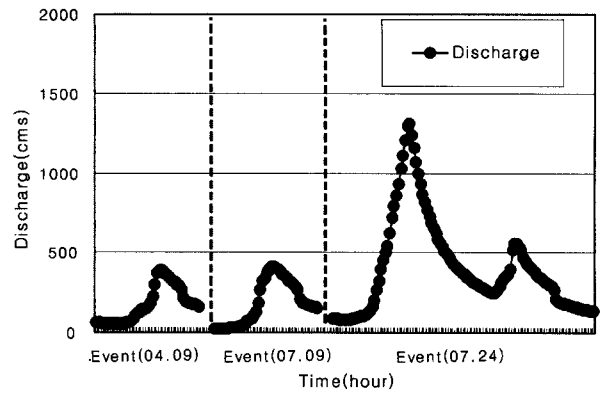


Fig. 6. Discharge of Pyongchang river.

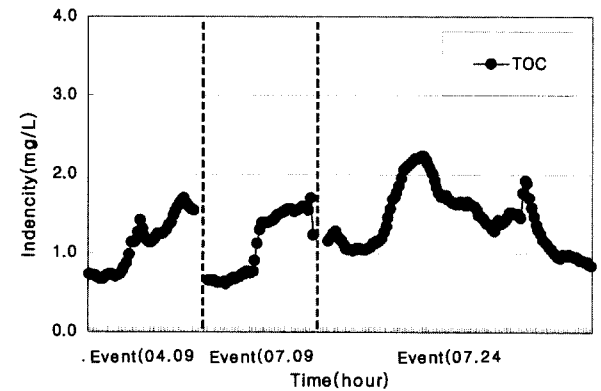


Fig. 7. TOC of Pyongchang river.

무강우시 보다 큰 것으로 분석되었으며 TOC는 유량 변동에 민감한 것으로 판단된다.

평창강 유역에서 동일기간동안 측정된 강우시 유량 및 TOC를 Fig 6과 7에 나타내었다.

Fig. 6과 7에서 보는 바와 같이 유량증가에 따라 수질변화가 매우 민감하게 반응하는 것으로 나타났다.

평창강 유역은 도달시간이 짧은 내사천 유역에 비해 유로연장 및 유역면적이 커서 강우 발생시간으로부터 약 17시간에서 20시간 후에 유량증가 및 수질변화가 일어났다.

3.2. 유량 및 수질항목간 관계분석

유량과 수질간의 상관관계를 파악하기 위하여 내사천과 평창강 유역에서 취득된 유량, 수질 시자료를 이용하여 상관분석을 실시하였다. 내사천 유역에서는 강우시 유량과 수질의 상관성은 낮은 것으로 나타났으며 유역의 하상경사가 심하고 유량의 변화폭이 매우 커서 홍수시 유량계측에 다소 어려움이 있었다. BOD의 경우 일부 측정된 COD와 높은 상관관계를 보였다.

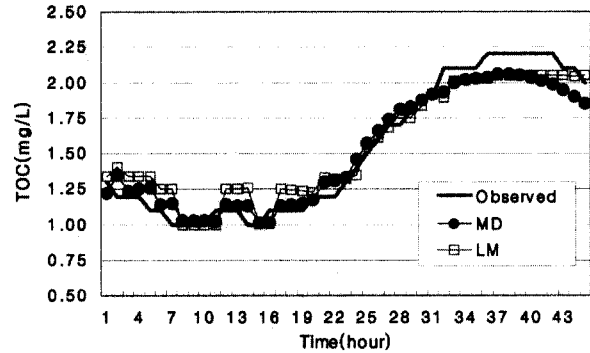
평창강 유역에서 1월, 4월, 7월의 측정기간별 유량과 TOC의 상관관계를 Table 5에 나타내었다. 계절별 상관관계를 보면 여름철인 7월에 0.84로 봄철인 4월의 0.68과 겨울철인 1월의 0.16에 비해 유량과 TOC의 상관관계가 크게 나타났다. 또한 유량과 TOC의 상관관계는 4월과 7월 모두 강우시에 큰 상관관계를 보였다.

**Table 5.** Correlation coefficient of discharge & TOC at Pyongchang river

Month	Rainy	Non-rainy	Total
1	-	0.16	0.16
4	0.62	0.43	0.68
7	0.80	0.23	0.84

3.3. 신경망 모형 적용

강우시 발생한 유량을 바탕으로 수질 변화 양상을 해석하기 위하여 비선형에 적합한 신경망 모형을 적용하였다. 내사천 유역에서는 2002년 7월 5일에 발생한 유량과 측정된 수질 자료를 학습하였으며 2002년 10월 5일에 측정된 자료를 검증 자료로 사용하였다. 평창강 유역에서는 2003년 4월 9일과 7월 9일에 발생한 유량자료와 수질자동측정망 자료를 학습하였으며 2003년 7월 24일에 측정된 자료를



**Fig. 9.** Comparison of the output result. (Pyongchang river)

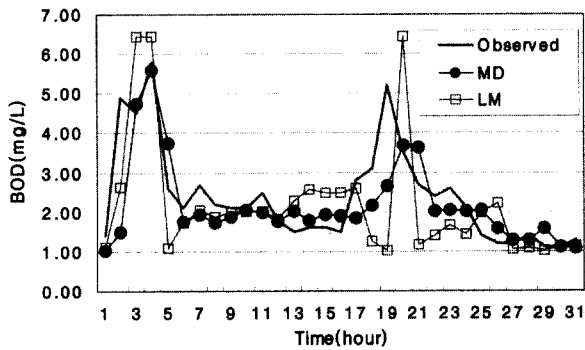
검증자료로 사용하였다.

내사천 유역에서 수행된 BOD, TN, TP에 대한 신경망 모형의 적용결과를 Fig. 8에 나타내었으며 평창강 유역에서의 TOC에 대한 적용결과를 Fig. 9에 나타내었다.

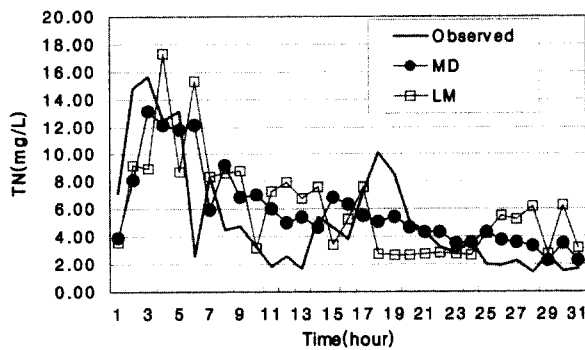
신경망 모형에 의한 수질예측 결과 내사천 유역에서는 다소 큰 오차를 보이고 있으나 수질변화가 양상은 파악되고 있는 것으로 나타났다. 하폭이 매우 작은 소하천인 내사천 유역에서 신경망 모형의 적용 결과가 다소 미흡한 원인은 취득된 자료의 양이 많지 않고, 측정된 유량과 수질의 상관성이 낮았기 때문으로 판단된다.

평창강 유역에서는 강우시의 수질변화에 대해 신경망 모형이 잘 적응하는 것으로 나타났으며 실측치에 근사하는 좋은 결과를 보였다. 평창강 유역에서는 유량과 수질의 변화가 서서히 진행되었고 안정적으로 측정되었으며 측정된 자료의 양이 많아서 충분한 검증이 이루어진 것으로 판단된다.

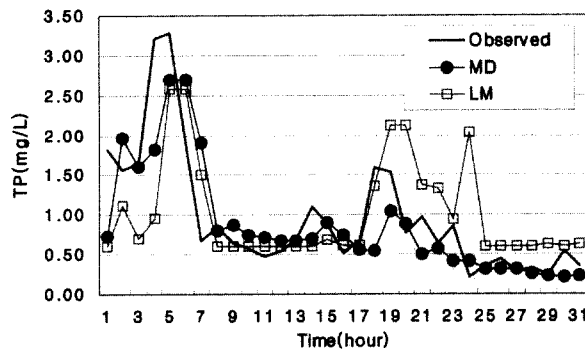
적용결과에 대한 분석 내용을 Table 6~7에 나타내었다. 내사천 유역에 적용된 LM, MD 신경망 모형에 의한 계산 값을 BOD, TN, TP의 실측치와 비교하였을 때 평균값에는



(a) BOD



(b) TN



(c) TP

**Fig. 8.** Comparison of the output result. (Naesa)

**Table 6.** Statistical analysis of output (Naesa)

	Model	Average (mg/L)	CC	RMSE (mg/L)
BOD	Observed	2.39	-	-
	MD	2.17	0.71	0.91
	LM	2.18	0.57	1.29
TN	Observed	5.34	-	-
	MD	5.94	0.64	3.15
	LM	6.15	0.37	4.36
TP	Observed	0.98	-	-
	MD	0.87	0.76	0.52
	LM	1.03	0.46	0.73

**Table 7.** Statistical analysis of output (Pyongchang river)

	Model	Average (mg/L)	CC	RMSE (mg/L)
TOC	Observed	1.54	-	-
	MD	1.53	0.99	0.09
	LM	1.52	0.99	0.10

근사하였으나 상관계수는 다소 낮았다. 두 신경망 모형을 비교하면 정성적 자료인 강우발생시간을 추가하여 학습한 MD 신경망 모형이 낮은 RMSE 값을 보였으며 상관성도 높은 것으로 나타났다.

평창강 유역에서는 TOC 모의를 위해 적용된 두 신경망 모형 모두 평균값에 근사하고 상관관계가 우수한 것으로 나타났다. RMSE 또한 LM 모형이 0.10 mg/L, MD 모형이 0.09 mg/L로 좋은 결과를 보였다.

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

#### 4. 결론

소하천인 내사천과 지방하천인 평창강을 대상으로 측정된 수질자료를 강우시와 무강우시로 구분하여 분석하고 강우시의 수질변화를 신경망 모형으로 예측하여 다음과 같은 결론을 얻었다.

1. 강우시에 측정된 수질이 무강우시에 비해 큰 평균값과 최대값을 보였으며 최소값과 최대값의 차이 또한 커서 표준편차도 큰 것으로 분석되었다.
2. 초기강우에 의해 발생된 유량이 유역의 수질변화에 큰 영향을 미치는 요인으로 나타났으며 평창강 유역에서 유량변동에 따른 TOC의 변화는 7월인 여름철에 가장 민감한 것으로 분석되었다.
3. 유량과 수질항목간의 관계해석을 위해 강우시에 적용한 신경망 모형은 소유역인 내사천 유역보다 유역면적이 큰 평창강 유역에서 적용성이 우수하였으며 정성적 자료인 강우발생시간을 활용한 모형이 좋은 결과를 보였다.

향후 수질자동측정망이 설치된 지점에서 실시간 유량관측시설을 동시에 운영하고 수심별, 횡단거리별 채취지점을

증가하여 양질의 자료를 취득하기위한 방안이 요구된다. 또한 강우에 따른 유량예측 시스템과 연계하여 수질변화를 미리 예측하고 이상 수질 발생시 빠르게 판단하여 수질사고 지점을 파악할 수 있는 시스템이 구축된다면 수질감시에 도움이 될 것으로 판단된다.

#### 사 사

본 연구의 일부는 환경부 지정 충북지역 환경기술개발센터의 연구비 지원으로 수행되었습니다.

#### 참고문헌

1. 김주식, 금강수계 수질자동측정망 추가 설치를 위한 우선순위 분석, 충남대학교 대학원 (2003).
2. 박석순, 이시진, 이용석, 최경식, 소수력 발전댐의 영향예측을 위한 한탄강 수질 모델 연구, *한국물환경학회지*, 15(1), pp. 89-98 (1999).
3. 안상진, 연규방, 연인성, 신경망 모형을 이용한 수질예측과 시스템 구축, *대한상하수도학회 한국물환경학회 공동추계학술발표회 논문집*, pp. 249-252 (2000).
4. 이상도, 중소하천의 유량과 수질변동 특성에 관한 연구, *홍익대학교 대학원* (1995).
5. 조재현, 목장지대 주변에 위치한 도암호의 수질관리, *상하수도학회지*, 12(3), pp. 41-47 (1998).
6. 조용진, 충주호 유역의 수질오염 발생특성 분석 및 관리방안 연구, *충북지역환경기술개발센터*, pp. 282-370 (2002).
7. 최지용, 비점오염원 유출부하량 조사방법, *국립환경연구원* (1994).
8. 한국수자원공사, *충주댐 유량측정 보고서* (2002).
9. Gerald, T. O., Water quality modeling for decision making, *Journal of water resources planning and management*, 118(3), pp. 295-307 (1992).
10. Ward, R., Loftis, J. and McBride, G., *Design of Water quality monitoring systems*, VNR (1990).