

연구논문

신경망과 퍼지 알고리즘을 이용한 하천 수질예측

이경훈 · 강일환* · 문병석** · 박진금***

전남대학교 토목공학과, 경호엔지니어링(주) 상하수도부*,
서남대학교 토목공학과**, 대한주택공사 광주전남본부***

(2004년 12월 17일 접수, 2005년 3월 10일 승인)

Water Quality Forecasting of River using Neural Network and Fuzzy Algorithm

Kyoung-Hoon Rhee · Il-Hwan Kang* · Byoung-Seok Moon** · Jin-geum Park***

Department of Civil Eng., Chonnam National University,
Department of Water and WasteWater, Kyong-ho Eng.*,
Department of Civil Eng., Seonam University**,
The Head Office of GwangJu Chonnam, Korea National Housing Corporation***
(Manuscript received 17 December 2004; accepted 10 March 2005)

Abstract

This study applied the Neural Network and Fuzzy theory to show water-purity control and preventive measure in water quality forecasting of the future river. This study picked out NAJU and HAMPYUNG as the subject of investigation and used monthly the water quality and the outflow data of KWANGJU2, NAJU, YOUNGSANNPO and HAMPYUNG from 1995 to 1999 to forecast BOD, COD, T-N, T-P water density. The datum from 1995 to 1999 are used for study and that of 2000 are used for verification. To develop model of water quality forecasting, firstly, this research formed Neural Network model and divided Neural Network model into two case - the case of considering lag and not considering. And this study selected optimal Neural Network model through changing the number of hidden layer based on input layer(n) from n to 3n. Through forecasting result, the case without considering lag showed more precise simulated result. Accordingly, this study intended to compare, analyse that Fuzzy model using the method without considering lag with Neural Network model. As a result, this study found

that the model without considering lag in Neural Network Network shows the most excellent outcome. Thus this study examined a forecasting accuracy, analyzed result and verified propriety through applying the method of water quality forecasting using Neural Network and Fuzzy Algorithms to the actual case.

Key words : Neural Network, Fuzzy, Lag Time, Input Layer, Hidden Layer

I. 서론

급속한 산업화와 인구의 도시집중 및 국민의 생활수준 향상으로 물 수요량의 급격한 증가와 수질오염은 심각한 사회문제로 대두되고 있다. 특히, 우리나라의 경우는 강우의 극심한 편중으로 인하여 6~9월의 우기가 지나면 10월에서 다음해 4~5월까지의 하천 유량이 줄어들어 하천의 자정능력과 희석 작용이 부족하여 오염현상이 심화되고 있다. 그리하여 수질오염의 방지 및 저감을 위하여 오·폐수 처리시설의 확대와 첨단화를 도모하여 환경 및 생태계 변화에 적극 대응하고 있는 실정이다.

다양해지는 오염원에 의한 수계의 수질오염을 효과적으로 관리하기 위하여 현재의 수질을 파악하고 장래 하천수질을 정확히 예측하는 방법이 요구되어지는 바, 본 연구에서는 수질변화 과정을 인간의 경험적 사고과정을 수학적으로 모형화한 신경망과 퍼지 이론을 적용하여 예측을 실시하였다.

신경망 이론을 이용한 하천의 수질모의에 관한 연구는 최근 활발히 이루어지고 있다. 안상진 등(2000)은 지점별 항목별 장·단기 수질농도예측을 위한 통합 시스템을 구축하여 최신의 자료에 맞도록 모형을 개선해 나가며 예측을 수행하였고 김만식 등(2001)은 학습 횟수 및 은닉층 그리고 노드 수에 변화를 주어 다층신경망 모형을 구축, 섬진강 유역의 쌍치와 관촌 지점의 수질을 모의하였다. 이들 연구와 최근의 수질모의에 관한 신경망 관련 연구는 변수에 지체시간을 고려하였다.

반면, 퍼지 이론을 이용한 수질모의 연구는 다

소 적으나 퍼지 클러스터링 알고리즘을 적용하여 전력부하를 예측한 연구 등 다양한 분야에 적용되고 있다.

본 논문에서는 장래 하천의 수질관리 및 방지대책을 제시하고자 수질 예측을 행함에 있어 신경망과 퍼지 모형을 각각 구성하였다. 우선 신경망은 지체를 고려한 경우와 고려치 않은 경우로 나누어 모형을 구성하였으며 두 방법 중 양호한 모의 결과를 가진 모형에 퍼지 이론을 적용하였다. 본 연구의 대상지점과 예측 수질항목은 영산강 수계의 분류인 나주와 함평의 BOD, COD, T-N, T-P수질농도이고, 대상자료는 광주2, 나주, 영산포, 함평 지점의 월별 수질항목이므로 지체시간을 고려한 방법과 고려치 않은 방법으로 각각 나눠 모형을 구성하였다. 또한, 입력층수(n)에 따른 은닉층의 개수를 n부터 3n까지 변화를 시키면서 최적의 신경망 모형을 선별하였다. 여기서, 모멘텀 상수와 초기 학습율은 모든 모형에서 국부최소값(local minimum)에 빠지지 않도록 각각 0.7과 0.9를 일괄적으로 사용하였다.

본 연구의 목적은 대상수계에서 유량변화 및 효율에 따른 수질농도예측을 하여 장래 하천의 수질관리 및 방지대책을 강구하는 기본적 자료를 제시하는데 필요한 예측방법의 모형을 제안하고 그 적용성을 검증하고자 하는 것이다.

II. 하천의 수질 예측모형

1. 신경망 모형

적응식 신경망의 기본적인 학습규칙은 경사하

며 광주2, 나주, 영산포, 함평교의 위치를 나타내었다. 본 연구의 대상지점은 나주와 함평이다.

2. 연구자료분석

본 연구에서는 독립변수를 동정하기 위해 단계선택법을 이용하였으며, 자기상관 및 상호상관분석을 통해 동정된 변수에 대한 지체시간(lag time)을 결정하여 모형을 구성하였다. 또한 한편으론 대상자료의 수가 적고 월별 자료이므로 지체시간을 주지 않고 모형을 구성하였다. 따라서 상관분석, 자기상관 및 상호상관의 분석 방법을 통해 얻어진 모형의 구성은 IV장에 제시하기로 한다.

IV. 모형의 구성

연구자료분석의 결과를 토대로 예측 대상지점인 나주와 함평 두 지점의 BOD, COD, T-N, T-P 수질농도 예측을 위한 모형은 표 1~4와 같다.

우선 신경망 이론을 적용한 모형은 지체를 고려한 경우와 지체를 고려하지 않은 모형으로 나누어 모형구성을 하였으며 통계분석 결과 지체를 고려하지 않은 모형이 우수한 예측결과를 보였으므로, 퍼지 이론을 적용한 모형은 지체를 고려하지 않은 모형으로 구성하였다. 따라서 신경망 모형구성에서 지체를 고려하지 않은 모형은 각각의 예측지점에 수질예측항목을 적어 모형을 표기하였고, 지체를 고려한 모형은 그 앞에 L을 적어 표기하였다. 한편으로 퍼지 모형구성에서 지체를

Table 1. Construction Model according to the water quality item of NAJU(Lag Time)

Model	Input Variable	Output Variable
LN BOD	N BOD _{t-1} , K COD _{t-1} , N COD _{t-1} , K T-N _{t-1} , N T-N _{t-1} , K DO _{t-1} , N DO _{t-1} , K SS _t , N SS _t , K pH _t , N pH _t , N Qt, N Tt	N BOD _t
LN COD	N COD _{t-1} , K BOD _{t-1} , N BOD _{t-1} , K SS _t , N SS _t , K T-N _{t-1} , N T-N _{t-1} , K T-P _{t-1} , N T-P _{t-1} , N Qt	N COD _t
LN T-N	N T-N _{t-1} , K BOD _{t-1} , N BOD _{t-1} , N Tt, K T-P _{t-1} , N T-P _{t-1} , K COD _{t-1} , N COD _{t-1} , K SS _t , N SS _t , N Qt	N T-N _t
LN T-P	N T-P _{t-1} , K T-N _{t-1} , N T-N _{t-1} , K COD _{t-1} , N COD _{t-1} , N Qt	N T-P _t

Table 2. Construction Model according to the water quality item of NAJU(Non Lag Time)

Model	Input Variable	Output Variable
N BOD	K COD _t , N COD _t , K T-N _t , N T-N _t , K DO _t , N DO _t , K SS _t , N SS _t , K pH _t , N pH _t , K Qt, N Qt, K T _t , N T _t	N BOD _t
N COD	K BOD _t , N BOD _t , K SS _t , N SS _t , K T-N _t , N T-N _t , K T-P _t , N T-P _t , K Qt, N Qt	N COD _t
N T-N	K BOD _t , N BOD _t , K T _t , N T _t , K T-P _t , T-P _t , K COD _t , N COD _t , K SS _t , N SS _t , K Qt, N Qt	N T-N _t
N T-P	K T-N _t , N T-N _t , K COD _t , N COD _t , K Qt, N Qt	N T-P _t

Table 3. Construction Model according to the water quality item of HAMPYOUNG(Lag Time)

Model	Input Variable	Output Variable
LH BOD	H BOD _{t-1} , K COD _{t-1} , N COD _{t-1} , Y COD _{t-1} , H COD _{t-1} , K T-N _{t-1} , N T-N _{t-1} , Y T-N _{t-1} , H T-N _{t-1} , K DO _{t-1} , N DO _{t-1} , Y DO _{t-1} , H DO _{t-1} , K SS _t , N SS _t , Y SS _t , H SS _t , K pH _t , N pH _t , Y pH _t , H pH _t , H Qt, H T _t	H BOD _t
LH COD	H COD _{t-1} , K BOD _{t-1} , N BOD _{t-1} , Y BOD _{t-1} , H BOD _{t-1} , K SS _t , N SS _t , Y SS _t , H SS _t , K T-N _{t-1} , N T-N _{t-1} , Y T-N _{t-1} , H T-N _{t-1} , K T-P _{t-1} , N T-P _{t-1} , Y T-P _{t-1} , H T-P _{t-1} , H Qt	H COD _t
LH T-N	H T-N _{t-1} , K BOD _{t-1} , N BOD _{t-1} , Y BOD _{t-1} , H BOD _{t-1} , H T _t , K T-P _{t-1} , N T-P _{t-1} , Y T-P _{t-1} , H T-P _{t-1} , K COD _{t-1} , N COD _{t-1} , Y COD _{t-1} , H COD _{t-1} , K SS _t , N SS _t , Y SS _t , H SS _t , H Qt	H T-N _t
LH T-P	H T-P _{t-1} , K T-N _{t-1} , N T-N _{t-1} , Y T-N _{t-1} , H T-N _{t-1} , K COD _{t-1} , N COD _{t-1} , Y COD _{t-1} , H COD _{t-1} , H Qt	H T-P _t

Table 4. Construction Model according to the water quality item of HAMPYOUNG(Non Lag Time)

Model	Input Variable	Output Variable
H BOD	K COD _t , N COD _t , Y COD _t , H COD _t , K T-N _t , N T-N _t , Y T-N _t , H T-N _t , K DO _t , N DO _t , Y DO _t , H DO _t , K SS _t , N SS _t , Y SS _t , H SS _t , K pH _t , N pH _t , Y pH _t , H pH _t , K Q _t , N Q _t , Y Q _t , H Q _t , K T _t , N T _t , Y T _t , H T _t	H BOD _t
H COD	K BOD _t , N OD _t , Y BOD _t , H BOD _t , K SS _t , N SS _t , Y SS _t , H SS _t , K T-N _t , N T-N _t , Y T-N _t , H T-N _t , K T-P _t , N T-P _t , Y T-P _t , H T-P _t , K Q _t , N Q _t , Y Q _t , H Q _t	H COD _t
H T-N	K BOD _t , N BOD _t , Y BOD _t , H BOD _t , K T _t , N T _t , Y T _t , H T _t , K T-P _{t-1} , N T-P _{t-1} , Y T-P _{t-1} , H T-P _{t-1} , K COD _{t-1} , N COD _{t-1} , Y COD _{t-1} , H COD _{t-1} , K SS _t , N SS _t , Y SS _t , H SS _t , K Q _t , N Q _t , Y Q _t , H Q _t	H T-N _t
H T-P	K T-N _t , N T-N _t , Y T-N _t , H T-N _t , K COD _t , N COD _t , Y COD _t , H COD _t , K Q _t , N Q _t , Y Q _t , H Q _t	H T-P _t

고려하지 않은 모형은 각각의 예측지점에 수질예측항목을 적어 그 앞에 F를 적어 표기하였다. 퍼지 모형구성은 신경망 모형의 지체를 고려하지 않은 경우와 동일하다. 입력변수 중 유량은 Q, 수온은 T로 표기하였다. 표 1~4에서 나주, 광주, 영산포, 합평지점은 영문의 첫 글자를 따서 N, K, Y, H로 표기하였다.

표 1과 표 2는 지체를 고려한 경우와 고려치 않은 경우의 나주지점의 수질농도 예측모형이다.

표 3과 표 4는 지체를 고려한 경우와 고려치 않은 경우의 합평지점의 수질농도 예측모형이다.

V. 분석방법 및 결과고찰

1. 분석방법

수질 예측모형의 적용타당성을 평가하기 위하

여 각 모형에서 계산된 결과를 오차 및 통계적인 검증방법을 이용하여 검토하였다. 본 연구에서는 시계열 자료에 대한 적용타당성 문제를 다룰 때 가장 일반적으로 사용되고 있는 평균제곱근오차 (root mean square error: RMSE), 절대평균오차 (mean absolute error: MAE) 및 상관계수 (correlation coefficient: CC)를 사용하였다.

2. 결과고찰

표 5는 나주와 합평지점의 수질농도 예측항목인 BOD, COD, T-N, T-P의 지체를 고려한 경우와 고려치 않은 경우의 신경망 예측모형을 통계 분석한 결과로 검증자료의 통계분석결과를 보면 지체를 고려치 않은 경우가 지체를 고려한 경우보다 양호한 모의 결과를 나타내고 있음을 알 수 있다. 이렇듯 지체를 고려하지 않은 모형이 양호

Table 5. Result of simulated by Model NAJU and HAMPYOUNG(Neural Network)

Model	Checking Data of NAJU(2000)				Model	Checking Data of HAMPYOUNG(2000)			
	MAE	RMSE	CC	Hidden Layer nodes		MAE	RMSE	CC	Hidden Layer nodes
LN BOD	0.878	1.061	0.868	3n	LH BOD	0.684	0.893	0.854	2.5n
N BOD	0.638	0.711	0.947	1.5n	H BOD	0.678	0.947	0.850	3n
LN COD	1.243	1.691	0.871	2n	LH COD	0.824	1.251	0.763	1.5n
N COD	1.198	1.495	0.887	2n	H COD	0.773	0.989	0.860	n
LN T-N	1.669	2.069	0.734	3n	LH T-N	1.576	1.798	0.763	3n
N T-N	1.459	1.835	0.794	2n	H T-N	0.917	1.212	0.901	2.5n
LN T-P	0.182	0.239	0.758	n	LH T-P	0.061	0.077	0.844	2.5n
N T-P	0.157	0.199	0.850	2.5n	H T-P	0.058	0.079	0.847	n

Table 6. Result of simulated by Model NAJU and HAMPYOUNG(Fuzzy)

Model	Checking Data of NAJU(2000)			Model	Checking Data of HAMPYOUNG(2000)		
	MAE	RMSE	CC		MAE	RMSE	CC
FN BOD	1.229	1.401	0.78	FH BOD	1.45	1.79	0.61
FN COD	2.113	2.66	0.73	FH COD	1.29	1.79	0.68
FN T-N	4.73	6.52	0.35	FH T-N	2.14	2.66	0.65
FN T-P	0.61	0.86	0.45	FH T-P	0.01	0.13	0.77

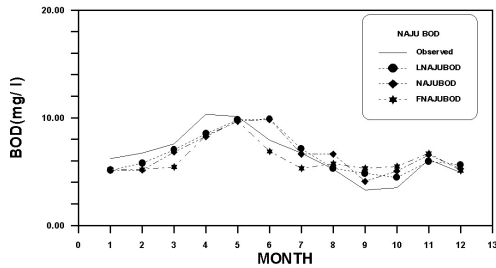


Fig. 2. Estimation of NAJU BOD

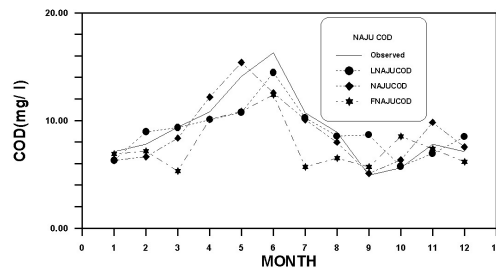


Fig. 3. Estimation of NAJU COD

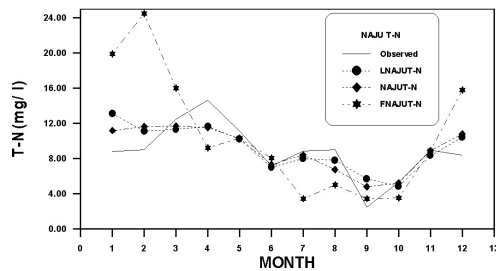


Fig. 4. Estimation of NAJU T-N

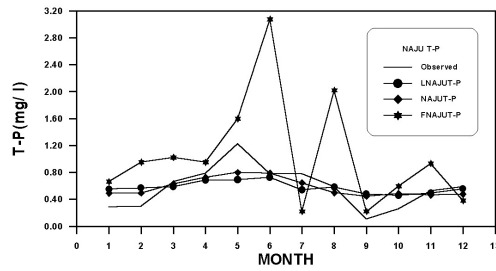


Fig. 5. Estimation of NAJU T-P

한 모의 결과를 나타내는 이유는 대상자료의 수가 적고 월별자료인 까닭으로 판단된다.

신경망 모형의 지점별 모의 결과를 살펴보면, 나주지점은 BOD, COD, T-P, T-N의 기술한 순으로 양호한 모의 결과를, 함평지점은 T-N, COD, BOD, T-P의 기술한 순으로 양호한 모의 결과를 보인다. 또한 각 지점별 모형의 적정한 은닉층 개수를 살펴보면 나주지점의 BOD, COD, T-N, T-P 수질농도 모형에선 각각 21, 20, 24, 15개임을 알 수 있으며, 함평지점의 BOD, COD, T-N, T-P 수질농도 모형에선 84, 20, 60, 12개임을 알 수 있다.

표 6은 나주와 함평지점의 수질농도 예측항목

인 BOD, COD, T-N, T-P의 지체를 고려치 않은 경우의 퍼지 예측모형을 통계분석한 결과로 검증 자료의 통계분석결과를 보면 나주지점에선 BOD, COD항목이, 함평지점에서 T-P항목이 양호한 예측결과를 보일 뿐 다른 항목은 저조한 예측결과를 보이고 있다. 또한, 신경망 모형이 퍼지 모형보다 우수한 예측결과를 보임을 알 수 있다.

그림 2~5는 신경망과 퍼지 모형의 지체를 고려치 않은 경우와 신경망 모형의 지체를 고려한 경우의 3가지 경우를 가지고 나주지점의 BOD, COD, T-N, T-P수질농도의 검증자료 예측결과를 도시한 것으로 수질농도의 변화가 심한 곳에서는

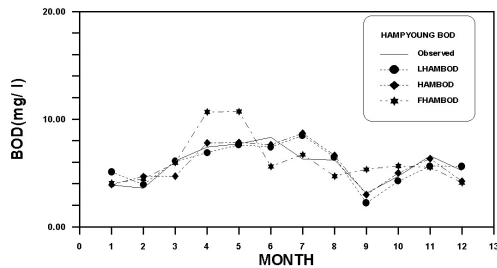


Fig. 6. Estimation of HAMPYOUNG BOD

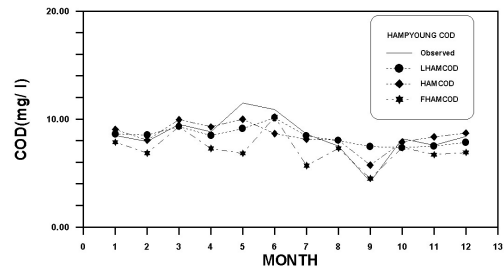


Fig. 7. Estimation of HAMPYOUNG COD

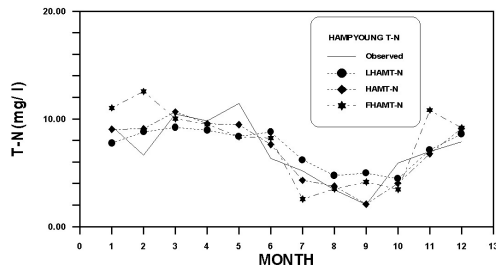


Fig. 8. Estimation of HAMPYOUNG T-N

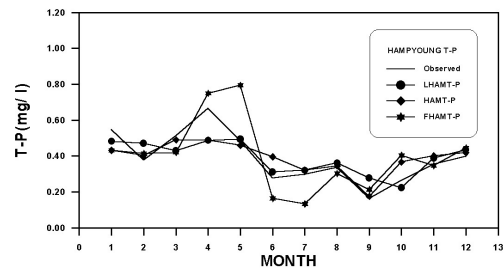


Fig. 9. Estimation of HAMPYOUNG T-P

예측치와 실측치간에 어느정도 차이가 있음을 보이고 있다. 또한 퍼지 모형의 T-N과 T-P수질농도는 저조한 모의 결과를 보인다. 3가지 모형 중에서 신경망의 지체를 고려치 않는 모형이 가장 양호한 모의 결과를 보임을 확인할 수 있다.

그림 6~9는 신경망과 퍼지 모형의 지체를 고려치 않은 경우와 신경망 모형의 지체를 고려한 경우의 3가지 경우를 가지고 함평지점의 BOD, COD, T-N, T-P수질농도의 검증자료 예측결과를 도시한 것으로 수질농도의 변화가 심한 곳에서는 예측치와 실측치간에 어느 정도 차이가 있음을 보이고 있다. 함평지점도 나주지점과 마찬가지로 신경망의 지체를 고려치 않는 모형이 다른 모형에 비해 비교적 양호한 모의 결과를 보이고 있다.

VI. 결론

본 연구에서는 영산강 본류의 수질 예측을 위한 활용가능성을 판단하기 위하여 지체를 고려한

경우 및 고려치 않은 경우의 신경망 모형, 지체를 고려치 않은 퍼지 모형을 구성하여 예측 정확도를 검토하고 결과를 분석하였다. 그 결과 다음과 같은 결론을 얻었다.

1. 수질 예측지점인 나주와 함평의 모든 수질농도 예측모형에서 지체를 고려치 않은 신경망 모형이 가장 양호한 모의결과를 나타냄을 알 수 있다.
2. 나주지점의 지체를 고려치 않은 신경망 모형은 BOD, COD, T-N, T-P의 기술 순으로 양호한 모형임을 알 수 있고 BOD모형의 검증자료 통계분석결과 MAE는 0.638, RMSE는 0.711, CC는 0.947의 예측결과로 예측치가 실측치를 잘 따라감을 알 수 있다.
3. 함평지점의 지체를 고려치 않은 신경망 모형은 T-N, COD, BOD, T-P의 기술 순으로 양호한 모형임을 알 수 있었고, T-N모형의 검증자료 통계분석결과 MAE는 0.917, RMSE는 1.212, CC는 0.901로 상당히 우수한 예측력을 보였다.
4. 신경망 모형이 퍼지 모형보다 더 적합한 모

형으로 판단되었으나 퍼지 모형에선 나주지점의 BOD, 합평지점의 T-P수질농도가 양호한 예측결과를 보였다.

본 연구에 사용된 자료가 월자료이므로 예측의 정확도가 다소 떨어지지만 다년간에 걸쳐서 축적된 양질의 시자료 또는 일자료가 구축된다면 좀 더 훌륭한 예측이 가능할 것이라고 생각된다.

참고문헌

- 김만식, 이요상, 심규철, 심순보, 2001, 신경망 모형을 이용한 하천의 수질예측 연구, 2001년도 학술발표논문집(II), 한국수자원학회, 925-930.
- 김이곤, 1993, 퍼지모형을 이용한 전력부하 예측에 관한 연구, 전남대학교대학원 박사학위논문.
- 노형진, 2000, Excel2000에 의한 통계적 조사방법, 형설출판사, 189-205, 303-315.
- 도용태, 김일근, 김종완, 박창현, 1997, 인공지능 개념 및 응용, 희중당, 217-246.
- 안상진, 연구방, 연인성, 2000, 신경망 모형을 이용한 수질예측과 시스템 구축, 2000년도 추계 학술발표회논문집, 대한상하수도학회·한국물환경학회, 249-252.
- 영산강 홍수통제소, 1995-2000, 한국수문조사연보, 건교부.
- 영산강 환경관리청, 1995-2000, 환경연보, 환경부.
- 오준성 등, 1996, 영산강 수질관리 종합대책 수립에 관한 연구(최종보고서), 전남대학교공업기술연구소, 광주광역시.
- Dan W. Patterson, 1996, Artificial neural Networks Theory and Application, Prentice Hall, 20-36.