

농업용 저수지의 수질 예측 모델을 위한 PSO (Particle Swarm Optimization) 알고리즘의 적용

권용수 · 배미정 · 황순진¹ · 박영석*

(경희대학교 생물학과, ¹건국대학교 환경과학과)

Application of Particle Swarm Optimization (PSO) for Prediction of Water Quality in Agricultural Reservoirs of Korea. Kwon, Yong-Su, Mi-Jung Bae, Soon-Jin Hwang¹ and Young-Seuk Park* (Department of Biology, Kyunghee University, Seoul 130-701, Korea; ¹Department of Environmental Science, Konkuk University, Seoul 143-701, Korea)

In this study, we applied a Particle Swarm Optimization (PSO) algorithm to predict the changes of chlorophyll-*a* related to environmental factors in agricultural reservoirs in Korean national scale. Data were obtained from water quality monitoring networks of reservoirs operated by the Ministry of Agriculture and Forestry and the Ministry of Environment of Korea. From the database of the monitoring networks, 290 reservoirs were chosen with variables such as chlorophyll-*a* and 13 environmental factors (COD, TN, TP, Altitude, Bank height, etc.) measured in 2002. Based on Carlson's trophic status index, reservoirs were divided into five groups, and most agricultural reservoirs (TSI_{CHL} 64.1%, TSI_{TP} 75.5%) were in the eutrophic states. The groups were discriminated with environmental variables, showing that COD, DO, and TP were important factors to determine the trophic states. MLP-PSO (Multilayer perceptron (MLP) with PSO for the optimization) was applied for the prediction of chlorophyll-*a* with environment factors, and showed high predictability ($r=0.83$, $p < 0.001$). Additionally, the sensitivity analysis of the MLP-PSO model showed that COD had the strongest positive effects on the concentration of chlorophyll-*a*, and followed by TP, TN, DO, whereas altitude and bank height had negative effects on the concentration of chlorophyll-*a*.

Key words : agricultural reservoir, water quality, trophic status index, particle swarm optimization, discriminant analysis

서 론

우리나라 대부분의 저수지는 여름철 부영양화로 인해 녹조현상이 빈번하게 일어나고 있다(농림부, 2000). Rawson (1995)에 따르면 호소의 깊이가 얕을수록 부영양화가 쉽게 발생하는데, 우리나라 저수지의 대부분이 수심이 얕고 노후화되어 있기 때문에(Hwang *et al.*, 2003)

부영양화가 쉽게 되며, 저수지의 노후화로 인해 침전물로부터 영양염류의 용출 가능성이 증가하게 되어 부영양화의 가능성이 높아지는 것을 의미한다.

Harper (1992)에 따르면 저수지 내 제한 영양염인 인의 유입 증가는 저수지 부영양화를 발생시킨다. 저수지의 부영양화는 저수지 내 식물플랑크톤을 증가시켜 저수지 투명도를 감소시키고, 저층의 식물플랑크톤 부패로 용존산소가 감소하는 등의 수질 변화를 초래한다(신 등,

* Corresponding author: Tel: 02) 961-0946, Fax: 02) 961-0244, E-mail: parkys@khu.ac.kr

2002). 또한 저수지의 부영양화는 저수지 내 투명도, 총 인 농도, chlorophyll-*a* (Chl-*a*) 농도와 밀접한 관계를 가지고 있으며, 이들 환경 인자는 저수지의 영양상태를 분류하는 기준으로 사용된다(Carlson, 1977). 특히 Chl-*a* 농도가 저수지 영양 상태를 나타내는 데 있어 가장 중요한 역할을 한다(Lu, 1999). 그러나 우리나라 농업용 저수지 수질 기준에는 Chl-*a* 농도에 대한 규제가 없어 부영양화 정도를 판단하고, 이를 근거로 수질개선 사업을 시행할 법적 근거가 없다(이 등, 2003).

농업용저수지의 수질을 관리하기 위해서는 지속적인 모니터링이 필요하다. 그러나 우리나라에 있는 모든 저수지를 효율적으로 모니터링하고 관리하기란 현실적으로 매우 어려운 일이다. 따라서 이들 저수지들의 이화학적, 수리수문학적 특성을 분석하고, 그 특성에 따른 농업용저수지의 수질 예측관리가 필요하다. 이 등(2007)은 농업용저수지 내에서 수질변이의 특성을 파악하기 위해 영양염류와 조류 생산력의 관계를 분석하였으며, 전 등(2002)은 우리나라 농업용 저수지에서 물리학적 인자의 영향에 대한 연구를 하였다. 그러나, 기존의 수질 예측 모델들에서는 저수지의 이화학적 특성들에 초점이 맞춰져 상대적으로 수리수문학적 특성들은 그 중요도가 낮았다. 따라서 본 연구에서는 우리나라 농업용 저수지에서 저수지의 이화학적, 수리수문학적 환경인자 특성에 따라 저수지 수질을 예측하고, 이를 통해 효율적인 농업용 저수지 수질 관리에 기여하고자 한다.

재료 및 방법

1. 저수지 자료

농업용 저수지의 수질을 예측하기 위해 1990년부터 농림부와 환경부에서 실시하는 전국수질측정망 자료의 일부를 이용하였다. 총 670개의 저수지 중 대규모 호소를 배제하기 위하여 유효 저수량이 2천만 m³ 이하인 저수지만을 고려하였으며, 이용 가능한 자료를 고려하여 290개의 저수지를 선정하였다. 그리고 비교적 자료의 균질성이 높은 2002년도의 자료 중 녹조발생이 높은 9월 전후에 조사된 값을 선택하였다. 선정된 290개 저수지에서 chlorophyll-*a* (Chl-*a*) 농도와 13개 이화학적, 수리수문학적 환경자료를 선택하였다. 환경인자는 결측자료의 비율이 적은 6개의 이화학적 수질 인자인 용존산소량(DO), 화학적 산소요구량(COD), 총 인(TP), 총 질소(TN), 부유물질(SS), 온도(Temperature)를 선택하였다(Fig. 1). 그리고 저수지의 수리수문학적 특성을 반영하기 위해 유역

면적(Catchment area), 수면적(Surface area), 제방 길이(Bank length), 제방 높이(Bank height) 등은 통계자료를 이용하였으며, 지리적·형태적 환경인자로 고도(Altitude), 저수지 길이(Reservoir length), 저수지 둘레길이(Circumference)는 전자지도를 이용하여 측정하였다(Fig. 1). Chl-*a* 농도는 0.1~173.4 mg m⁻³의 분포 범위를 보였으며, 이러한 Chl-*a* 변이에 따라 COD, TP, TN, SS 등이 유사한 분포 양상을 보였다. 그러나 고도 및 제방높이는 Chl-*a* 농도분포와 반대되는 농도구배를 보였다. 또한 저수지 둘레길이, 수면적, 저수지 길이 간에는 유사한 분포 양상을 보였다.

저수지 영양상태를 평가하기 위하여 Chl-*a* 농도와 TP 농도에 따른 Carlson(1977) 부영양화 지수(TSI)를 식(1)과 (2)에 따라 계산하였다.

$$TSI_{CHL} = 9.81 \times \ln(\text{Chl-}a) + 30.6 \quad (1)$$

$$TSI_{TP} = 14.42 \times \ln(\text{TP}) + 4.15 \quad (2)$$

이때 Carlson(1977) 부영양화 지수(TSI)는 저수지 영양상태를 평가하는 가장 일반적인 방법으로, Chl-*a* 및 TP의 농도 단위는 $\mu\text{g L}^{-1}$ 이다. 계산된 TSI를 이용하여 Kratzer와 Brenzarick(1981)의 영양상태 판정표에 근거하여 분석대상 저수지의 영양상태를 등급에 따라 5개 그룹으로 구분하였다. 이때 각 영양단계에 속하지 않는 TSI값은 근사값으로 가장 근접한 영양단계로 반영하였다.

Chl-*a* 농도와 환경인자들 간의 연관성을 찾기 위하여 비모수 상관계수(Spearman rank correlation coefficient)를 계산하였으며, 통계 분석은 STATISTICA (StatSoft, 2004) 프로그램을 사용하였다.

2. 모델링 방법

분석대상 농업용 저수지를 TSI 값에 따라 5개 그룹으로 구분하였으며, 이를 환경인자를 사용하여 판별분석을 하였다. 그리고 MLP-PSO (Multilayer perceptron-Particle swarm optimization) 모델을 이용하여 저수지의 환경인자 특성에 따른 Chl-*a* 농도를 예측하였다. 분석자료 중 Chl-*a*, TP, TN, SS, 고도, 제방 길이, 저수지 길이, 저수지 수면적, 저수지 둘레길이는 변이 폭이 크기 때문에 이를 줄이기 위하여 log 변환을 하였다. 이때 log 값이 0이 되는 것을 방지하기 위하여 각 변수의 값에 1을 더해 계산하였다. 또한 각 변수들의 중요도를 균등하게 고려하기 위하여 모든 변수는 최소-최대 변환식(식(3))을 사용하여 각 변수의 범위를 0과 1 사이의 값으로 변환시켰다.

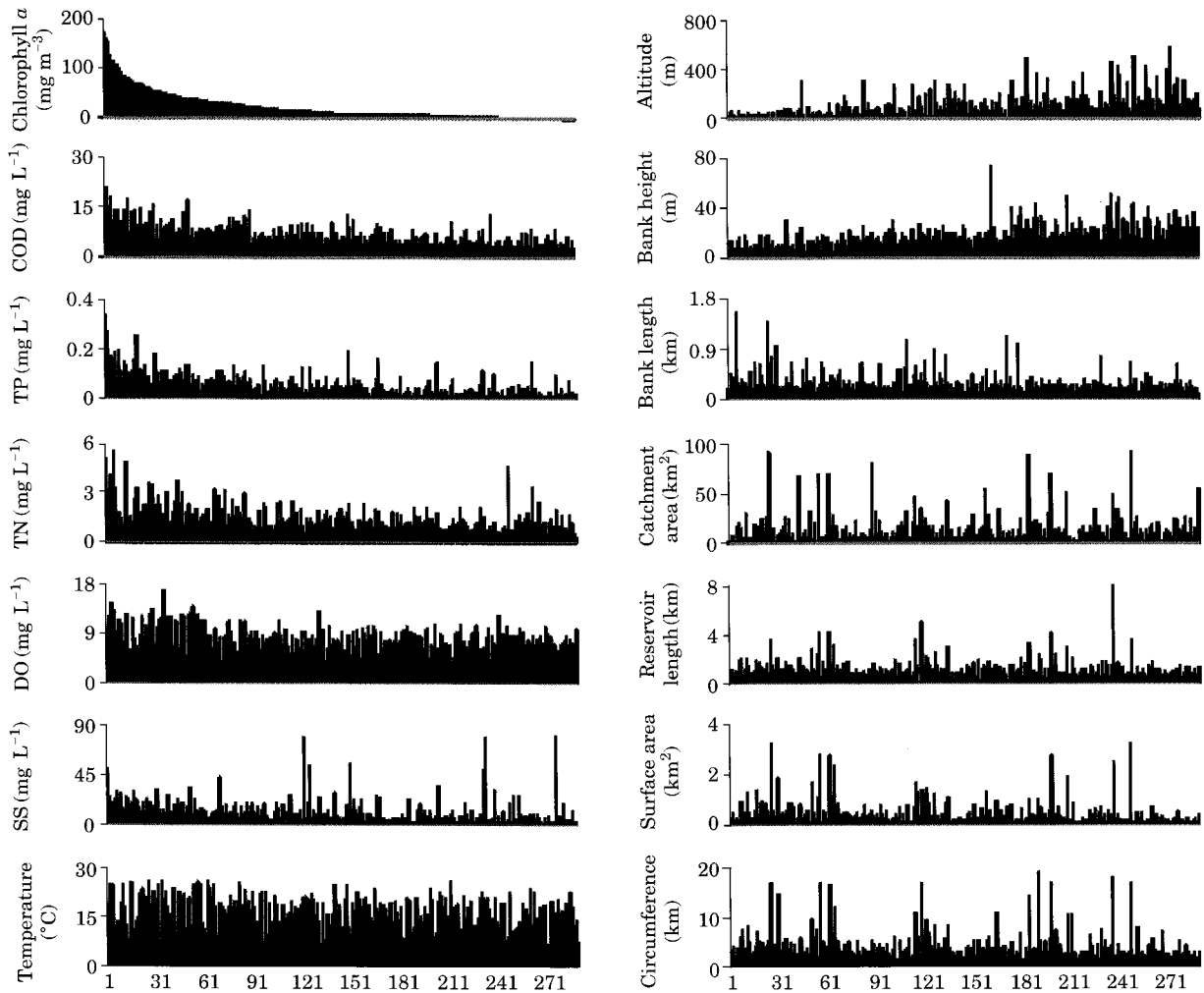


Fig. 1. Differences of environmental variables according to changes of chlorophyll-a concentration in 290 agricultural reservoirs, Korea in 2002.

$$b_i(x) = \frac{(x_i - \min(x))}{\max(x) - \min(x)} \quad (3)$$

여기서 $b_i(x)$ 는 표본 i 의 변수 x 의 변환된 값, $\min(x)$ 및 $\max(x)$ 는 각각 변수 x 의 최소값 및 최대값을 나타낸다.

1) 판별분석

판별분석은 하나 이상의 판별 변수의 조합으로 이루어진 판별 함수에 의해 그룹 내 분산에 대해 그룹 간 분산을 최대화하고자 하는 데 목적을 둔다(Wilson, 2002). TSI_{CHL} 값을 이용해 우리나라의 농업용 저수지의 영양상태를 분류한 결과를 바탕으로 판별분석에 필요한 그룹 변수를 지정하고, 측정 변수인 12개 환경인자들을 통해 개체의 그룹을 판별한다. 그 결과는 정준 분석을 사용하여 분석하였다. 판별분석을 위해서 통계분석 패키지인

STATISTICA (StatSoft, 2004)를 사용하였다.

2) Particle Swarm Optimization (PSO)

PSO는 컴퓨터 지능 분야에서 각광받고 있는 최신 기법으로(Birge, 2003), 새나 물고기의 사회적 학습 행동을 나타내기 위해 Kennedy and Eberhart (1995)가 고안한 확률적 최적화 기법에 바탕을 두고 있다. PSO 알고리즘의 원리는 각 입자는 최적의 해를 얻기 위해 다차원 탐색공간을 날아다니며, 그들 자신과 그들 이웃의 경험에 대한 정보를 이용하여 최적의 위치로 이동해 간다(Chau, 2006). 즉, 무리가 최적의 방향을 찾아가는 과정에서 무리 내부의 입자 방향이 지금까지의 자신의 경험과 무리 전체에 공유되어 있는 정보를 기초로 하여 행동한다는 개념을 최적화 과정에 도입한 것이다.

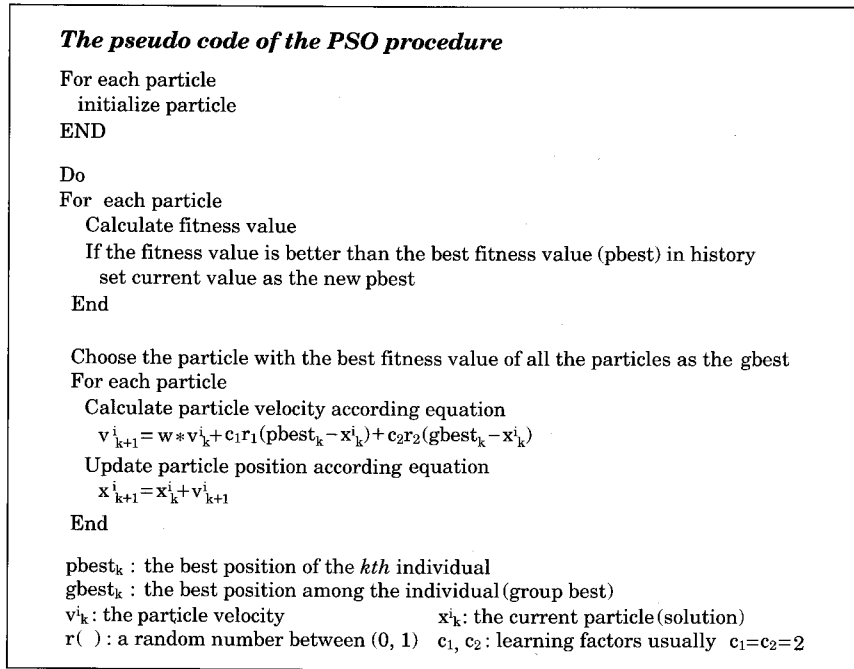


Fig. 2. The pseudo code of the PSO.

PSO 알고리즘은 다음과 같은 연산과정으로 이루어진다(Fig. 2). 먼저 입자들의 초기 위치와 속도를 최적화하기 위해 d 차원의 해 공간 상에서 입자들의 위치와 속도를 임의로 설정한다. 이후 현재 단계의 위치, 속도 정보와 목적 함수값을 이용하여 $k+1$ 번째의 속도 벡터를 식(4)에 따라 조정한다(Shi and Eberhart, 1998).

$$v_{k+1}^i = \omega \times v_k^i + c_1 \times r_1 (pbest_k - x_k^i) + c_2 \times r_2 (gbest_k - x_k^i) \quad (4)$$

이때 r_1 과 r_2 는 난수로 0과 1 사이의 값을 가지며, ω 는 관성 가중치, c_1 과 c_2 는 가속상수로 2이다. 그리고 v_k^i 는 i 번째 입자의 k 번째 속도, x_k^i 는 i 번째 입자의 k 번째 위치이다. $pbest_k$ 는 i 번째 입자의 현재까지의 최적 위치, $gbest_k$ 는 군집의 최적위치이다. 다음으로 i 번째 입자의 $k+1$ 번째 위치는 앞서 구한 속도 벡터를 이용하여 새로운 값으로 갱신되며 이때 식(5)를 사용한다.

$$x_{k+1}^i = x_k^i + v_{k+1}^i \quad (5)$$

본 연구에서는 PSO 알고리즘을 지도학습 인공신경회로망의 일종인 다층신경회로망(MLP, multilayer perceptron)의 최적화에 적용하였다(Fig. 3). 모델의 입력층에는 13개 환경인자 값을, 출력층에는 입력층의 환경인자에 대응되는 저수지 수질을 나타내는 Chl- a 농도 값을 제공하여 예측하였다. 즉, 3개의 층(입력층, 은닉층, 출력층)으로 구성된 MLP-PSO는 입력층에 환경인자들의 값을 할

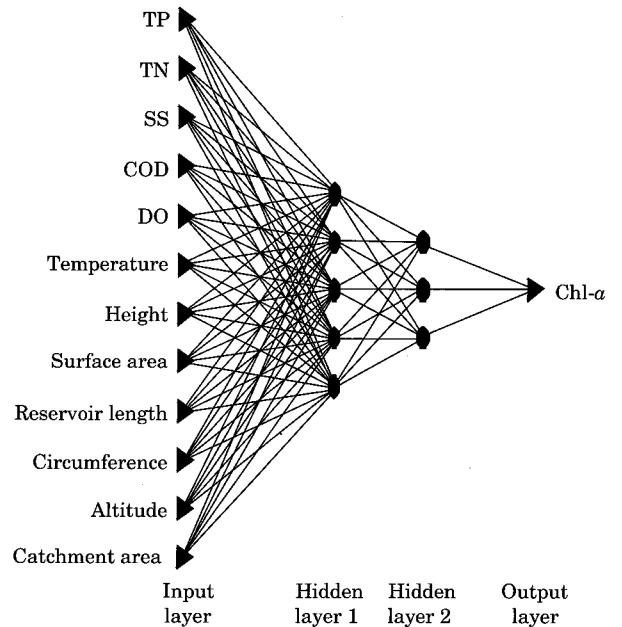


Fig. 3. Diagram of the MLP-PSO.

당하고 은닉층을 통해 최적해를 찾는다. 최적해는 출력층에 할당되는 Chl- a 의 값과 비교해 최량해를 도출하게 되는 것이다.

PSO 알고리즘을 적용한 MLP-PSO 모델은 입력층을

Table 1. The discrimination of 290 agricultural reservoirs into five different categories of trophic states using discriminant function analysis (Rows: Observed classifications, Columns: Predicted classifications).

		Predicted					Correct prediction (%)
		Ultra-oligotrophy	Oligotrophy	Mesotrophy	Eutrophy	Hyper-eutrophy	
Observed	Ultra-oligotrophy	0	0	1	1	0	0.0
	Oligotrophy	0	2	13	13	1	6.9
	Mesotrophy	0	0	31	40	2	42.5
	Eutrophy	0	2	12	96	9	80.7
	Hyper-eutrophy	0	0	0	23	44	65.7
	Total	0	4	57	173	56	59.8

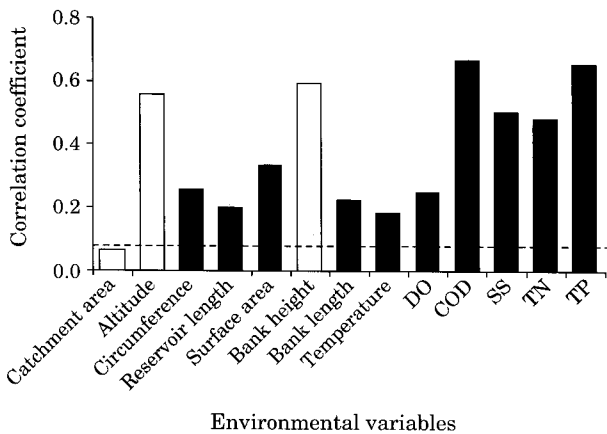


Fig. 4. Spearman rank correlation coefficients between chlorophyll-*a* and 13 environmental variables in the study reservoirs. The dotted line indicates statistically significant level ($p=0.05$) (dark bar: positive correlation coefficient, white bar: negative correlation coefficient).

서로 다른 조건으로 바꾸어 최적의 결과를 찾고자 하였다. 각 조건은 입력층에 사용된 환경인자 수를 달리한 것으로, 전체 13개 환경인자를 모두 사용한 경우, Chl-*a*와 유의한 상관관계를 보인 12개 환경인자만을 사용한 경우, 이화학적 수질 인자 6개만을 사용한 경우로 구성하였다. 각 환경인자 수에 따른 예측 결과는 은닉층의 수와 뉴런의 수를 변화시켜 Gbest (global best value)와 MLP-PSO 모델을 통해 얻은 예측값과 관찰값의 상관계수를 통해 나타내었다. 전체 290개 저수지 자료 중 194(전체의 2/3)개 저수지 자료는 모델의 학습에 사용하였으며, 학습에 사용되지 않은 나머지 96(전체의 1/3)개 저수지 자료는 모델의 검증에 사용하였다. MLP-PSO 모델은 Matlab (The Mathworks, 2001) 환경에서 운용되는 PSO toolbox (Birge, 2003)를 사용하였다.

결 과

1. 저수지 자료 특성

조사자료에서 Chl-*a* 농도는 이화학적 수질인자인 COD 및 TP와 높은 양의 상관관계를 보였으나(각각 $r=0.66$, $r=0.67$, $p<0.05$), 제방 높이 및 고도와는 음의 상관관계를 보였다(각각 $r=-0.60$, $r=-0.56$, $p<0.05$) (Fig. 4). 호소의 영양상태를 나타내는 지표 TSI_{CHL} 를 이용하여 전국 290개 농업용 저수지의 영양상태를 평가한 결과, 중영양 이하의 영양상태를 나타낸 저수지는 전체 290개 저수지 중 35.8% (104개)로 나타난 반면, 부영양 상태를 나타낸 저수지는 41% (119개), 과영양 상태는 23.1% (67개)를 나타내 전체 저수지의 64% 이상이 부영양화 되었다는 것을 보여주었다 (Fig. 5a). 또한 TSI_{TP} 를 통한 농업용 저수지의 영양상태를 평가한 결과에서는 중영양 상태 이하의 저수지는 전체 290개 저수지의 24.5% (71개)를 나타내었으며, 부영양 상태 이상의 저수지는 75.5% (219개)로 나타났다 (Fig. 5b).

2. 부영양화 정도의 판별분석

농업용 저수지의 부영양화 정도를 5개 그룹으로 나누고, 이를 환경인자 특성을 기준으로 판별분석 하였다. 그 결과 전체 290개 저수지 중 148개가 올바르게 판별되어 전체 판별적중률은 약 60%를 보였다 (Table 1). 각 그룹별 판별적중률은 부영양 및 과부영양 상태가 각각 80.7%와 65.7%로 높은 예측력을 보였다. 그러나 중영양 및 빈영양 상태에서는 각각 42.5% 및 6.9%로 낮게 나타났는데 이는 이들 그룹에서 많은 저수지들이 상위영양 상태로 예측되었기 때문이다. 판별분석 결과에 대하여 정준분석을 실시한 결과 첫 번째 축의 고유치는 1.176으로 총 변이의 91.1%를 설명하였으며, 두 번째 축의 고유치

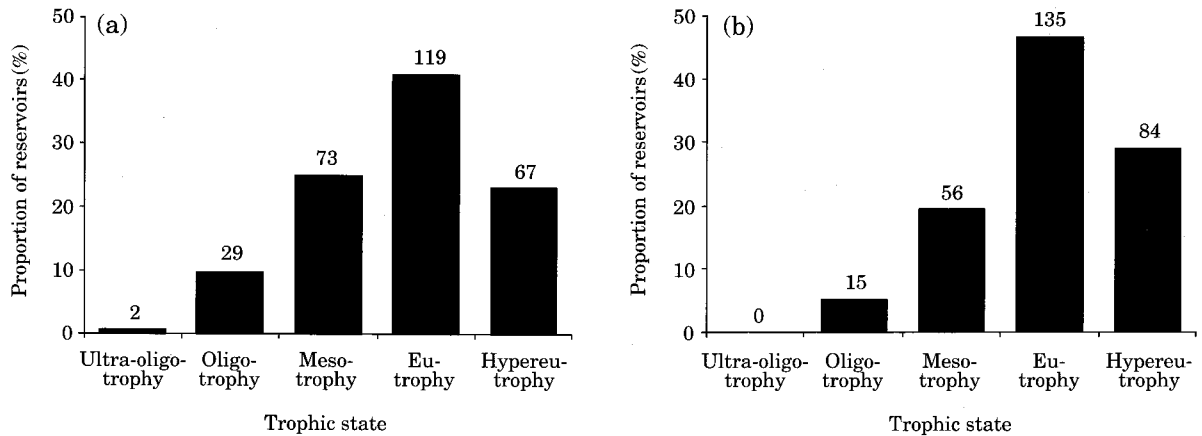


Fig. 5. Classification of 290 agricultural reservoirs based on the Carlson's trophic state index with chlorophyll-*a* (TSI_{CHL}) (a) and with total phosphorus (TSI_{TP}) (b).

Table 2. Standardized discriminant function coefficients through the discrimination analysis of the reservoir trophic status with 12 environmental variables.

Variable	Axis 1	Axis 2
Altitude	0.119	0.188
Circumference	0.241	0.436
Reservoir length	-0.154	-0.676
Surface area	-0.351	0.668
Bank height	0.211	0.656
Bank length	-0.025	-0.214
Temperature	-0.124	-0.065
DO	-0.446	0.446
COD	-0.516	-0.100
SS	0.285	-0.125
TN	-0.117	-0.062
TP	-0.395	0.556
Eigen value	1.176	0.059
Cum. proportion (%)	91.1	95.7

는 0.059로 총 변이의 4.6%를 설명하여, 첫 번째 축과 두 번째 축이 총 변이에 대해 96%를 설명하였다(χ^2 -test, $p < 0.05$) (Fig. 6). 정준분석에 따른 5개 그룹의 분포는 가장 오른쪽에 빈영양상태를 나타내는 그룹이 분포하며, 왼쪽으로 갈수록 부영양화가 진행되어 가장 오른쪽에 과부영양상태를 나타내는 그룹이 분포한다 (Fig. 6). 그러나 극 빈영양상태를 나타내는 그룹에 포함된 저수지의 수는 2개뿐이어서 명확한 경향성을 설명하기는 어렵다. 그룹의 판별분석에 영향을 주는 환경요인의 상대적 중요도를 정준 분석의 표준화 계수를 통해 나타내면 첫 번째 축에서 COD, DO, TP, 수면적 등의 순서로 상대적 중요성을 보

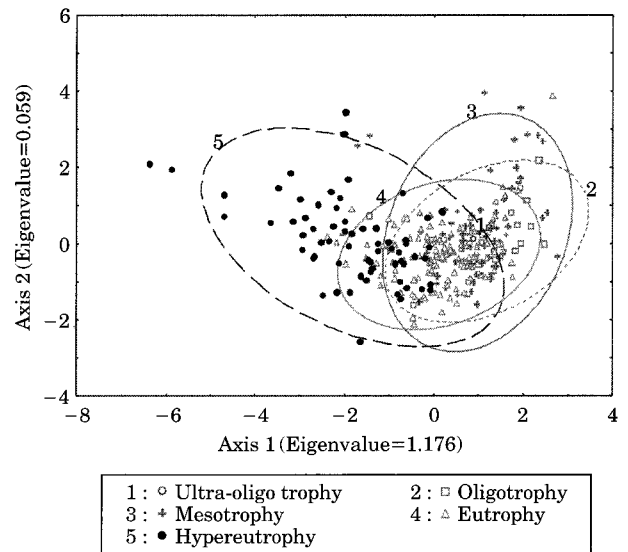


Fig. 6. Canonical analysis after discriminant function analysis for the discriminant analysis of the reservoir trophic status with 12 environmental variables.

였으며, 두 번째 축에서는 저수지 길이, 수면적, 제방길이, TP 등의 순서로 상대적 중요성이 나타났다 (Table 2). 또한 저수지의 고도 및 온도는 첫 번째 축과 두 번째 축에 대하여 상대적으로 낮은 기여도를 보였다.

3. Chl-*a* 농도 예측

MLP-PSO 모델을 이용하여 13개 환경인자 특성에 따라 Chl-*a* 농도를 예측하였다. MLP-PSO 모델은 2,000번

Table 3. Comparison of prediction results with different model structures of MLP-PSO.

Variables*	No. hidden layer(s) and neurons		Gbest	Correlation coefficient	
				Train	Validation
6	Single	3	0.021805	0.73005	0.83446
		5	0.020164	0.73282	0.83778
	Double	3, 5	0.025366	0.72972	0.84161
		5, 3	0.051263	0.73534	0.83836
12	Single	3	0.018207	0.75476	0.83071
		5	0.20548	0.65264	0.76512
	Double	3, 5	0.68402	0.48154	0.51231
		5, 3	0.082414	0.74159	0.80683
13	Single	3	0.023817	0.75206	0.83042
		5	0.29388	0.67247	0.75979
	Double	3, 5	0.40691	0.62292	0.67221
		5, 3	1.098	0.74888	0.82893

*6 variables: Temperature, DO, COD, SS, TN, TP; 12 variables: Altitude, Circumference, L-length, Surface area, Height, Length, Temperature, DO, COD, SS, TN, TP; 13 variables: Catchment area, Altitude, Circumference, L-length, Surface area, Height, Length, Temperature, DO, COD, SS, TN, TP

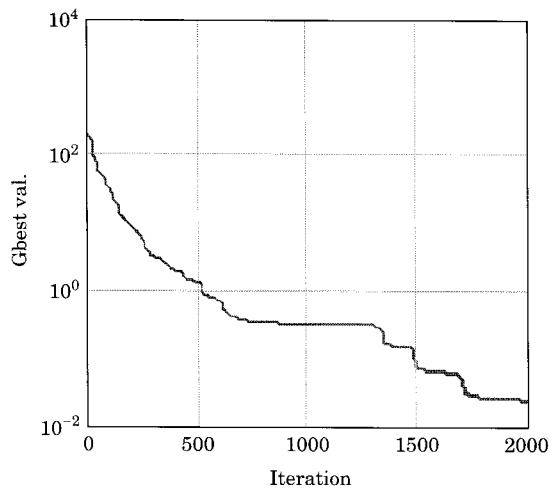


Fig. 7. An example of the changes of Gbest for the training of the MLP-PSO model.

학습반복을 통하여 수렴하였으며, 그 결과 각 조건에서 반복 횟수에 따라 Gbest 값이 감소하였다(Fig. 7). Gbest는 global error gradient로 현재의 값보다 양호한 Gbest 값이 얻어질 때만 새로 갱신되며, 0과 초기값 사이의 범위에서 형성된다.

각 조건에서의 MLP-PSO 모델 중 최적의 해를 가지는 모델을 선택하기 위해 Gbest 값과 예측력을 비교하였다. 세 조건 중 13개의 환경인자를 모두 사용한 경우에는 Gbest의 값이 뉴런의 수에 관계없이 전체적으로 높게 나타났다(Table 3). 반면 Chl-a 농도와 유의한 상관관계를

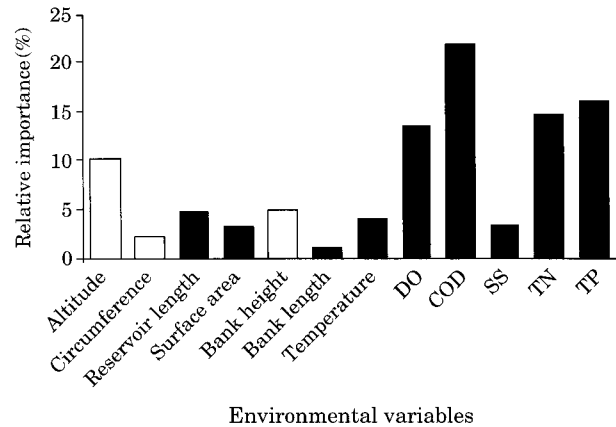


Fig. 8. Relative importance of 12 environmental variables for prediction of chlorophyll-a concentration in MLP-PSO model (dark bar: positive contribution, white bar: negative contribution).

보인 12개의 환경인자를 사용하는 경우와 6개의 환경인자만을 사용하는 경우 모두 13개 환경인자를 사용한 경우에 비해 낮은 Gbest 값을 보여 상대적으로 높은 정확성을 보였다. 특히 12개의 환경인자를 사용한 경우의 3개 뉴런으로 구성된 단일 은닉층에서 Gbest의 값이 0.018207로 가장 낮았으며, 모델을 통한 예측력 또한 다른 조건들에 비해 높게 나타났다(학습 $r=0.755$, 검증 $r=0.831$, $p < 0.05$). 이는 입력층에 12개의 뉴런을 사용하고, 단일 은닉층의 3개 뉴런으로 구성된 12-3-1 구조 모델이 가장 높은 효율성을 보였다.

저수지의 Chl-*a* 농도 결정에 영향을 미치는 환경인자들의 상대적 중요도를 평가하기 위하여 MLP-PSO 모델 중 가장 높은 예측력을 보인 12-3-1 구조 모델의 민감도 분석을 실시하였다. 그 결과 COD가 21.8%로 가장 높은 양의 기여도를 보였으며, 다음으로 이화학적 인자인 TP (16.1%), TN (14.6%), DO (13.6%) 순으로 높은 양의 기여도를 보였다 (Fig. 8). 반면 수리수문학적 인자인 고도 (10.1%) 및 제방높이 (4.9%)는 음의 기여도를 보였다.

고 찰

본 연구에서는 저수지의 이화학적, 수리수문학적 환경인자 특성에 따라 저수지의 영양상태를 판별분석하고, 또한 MLP-PSO 모델을 이용하여 저수지 수질을 나타내는 Chl-*a* 농도를 예측하였다. Hwang *et al.* (2003)에 따르면 우리나라 대부분의 농업용 저수지는 크기가 작고 수심이 10 m 이하로, 상대적으로 긴 생성시기 (>50년)를 가지고 있다. 이러한 농업용 저수지의 소형화 및 노후화로 인해 우리나라 대부분의 농업용 저수지는 부영양화가 되었거나 조류의 대 번성을 겪고 있다 (김 등, 2004). 본 연구에서도 이와 유사하게 Carlson (1977)의 부영양화 지수를 이용하여 우리나라 농업용 저수지의 영양상태를 평가한 결과, TSI_{TP}에서 부영양 상태 이상의 저수지 비율은 75.5%로 나타났으며, TSI_{CHL}에 따른 기준에서는 이보다 다소 낮은 64.1%로 나타났다 (Fig. 5). 부영양화 지수를 통해 식물플랑크톤 제한요인을 유추한 Havens (2000)의 연구에 따르면 TSI_{CHL} < TSI_{TP}인 경우에는 인이 아닌 다른 요인이 조류의 성장을 제한하는 것으로 유추할 수 있다. 이러한 결과는 농업용 저수지 내 식물성 플랑크톤에 축적된 인의 양이 포화상태를 이루어 공급된 인의 양보다 식물성플랑크톤의 1차 생산력에 의한 환원율이 감소한 것으로 판단된다 (Goldman *et al.*, 1987; 이 등, 2007).

저수지의 영양 상태를 5개 그룹으로 나누고 이를 환경인자들의 특성을 통해 판별분석한 결과, 전체 판별적중률은 약 60%를 보였다. 그 중 부영양화 이상의 상태는 70% 이상의 높은 적중률을 보였으나, 중영양 및 빈영양 상태에서는 보다 더 영양 상태가 높은 것으로 판별하여 적중률이 다소 낮게 나타났다. 이와 같이 보다 높은 영양 상태로 판정하는 것은 해당 저수지에서 Chl-*a*의 농도가 보다 더 높아질 가능성이 있다는 것을 반영해 준다. 판별 분석의 결과를 이용하여 정준분석을 수행한 결과, 첫 번째 축과 두 번째 축은 총변이에 대해 전체 96%의 설명력을 보여 부영양 상태가 명확히 구분될 수 있다는 것을

보였다 (Fig. 6). 특히 첫 번째 축의 고유치는 1.176으로 총 변이에 대해 91%의 설명력을 가지며, 표준화 계수에 의해 첫 번째 축에서의 변수들 간의 상대적 중요도는 COD, DO, TP의 순으로 나타났다 (Table 2).

MLP-PSO 모델에서 저수지의 Chl-*a* 농도는 환경특성에 따라 잘 예측되었다 (Table 3). MLP-PSO 모델은 MLP에 최적화기법의 일종인 PSO를 적용한 것으로, PSO는 유전자 알고리즘 (GA, genetic algorithm)과 같은 진화연산기법과 유사성을 지니고 있다. 이는 시스템이 초기에 임의의 해공간 (모집단)을 가지고, 새로운 발생을 통해 최적 조건을 찾는 것이다. 그러나 GA와는 달리 PSO는 교차나 변이와 같은 진화적 연산이 없이 잠재적 해집단인 입자가 전체 해집합의 공간을 탐색한다는 차이를 가진다 (Natarajan *et al.*, 2007). PSO의 장점은 상대적인 간결성, 연산의 효율성, 짧은 계산시간 안에 최적의 해를 생성할 수 있다는 것이다. 또한 PSO는 다른 확률적 방법에 비해 높은 수렴율과 높은 확률로 최적해에 도달한다 (Clerc *et al.*, 2002). 본 연구에서 MLP-PSO 모델의 결과 12개 환경인자를 사용한 12-3-1구조에서 가장 높은 예측력을 보였으며 (Table 3), MLP-PSO 모델의 민감도 분석을 통하여 저수지의 Chl-*a* 농도를 예측하는 데 있어 환경인자의 중요도를 평가한 결과 COD와 TP가 각각 21.8%와 16.1%로 가장 높은 양의 기여도를 보였으며, 고도 및 제방 높이는 음의 기여도를 보였다 (Fig. 8). 인자들 사이의 이러한 연관성은 우리나라 농업용 저수지를 이화학적 수질 특성에 따라 유형화하고 수리수문 및 지리학적 인자와 비교 분석한 Bae *et al.* (2007)의 결과와도 일치한다. 이는 호소 내에서 조류의 성장이 오염물질을 반영한다는 관점에서 볼 때, 조류의 성장은 무기물질에서 유기물질로의 전환을 의미하는 것으로 COD 농도와 Chl-*a* 농도와의 높은 상관성을 의미한다 (이 등, 2003). 또한 TP 농도는 부영양화를 발생시키는 남조류의 성장에 있어 가장 중요한 요소로 작용한다는 Smith *et al.* (1987)의 보고와도 일치한다. 한편, 본 연구에서 저수지 영양 상태를 평가를 위한 판별분석 결과 화학적 인자들의 상대적 중요성이 높게 나타났고, 고도 등의 인자들은 상대적으로 낮은 중요도를 보였다. 그러나 Chl-*a* 예측 모델의 환경인자 중요도 평가 결과, 이화학적 인자뿐만 아니라 고도 등의 지형학적 인자들의 중요성이 높게 나타났다 (Fig. 8). 즉, 고도가 높고, 제방 높이가 높은 지역은 산간지역으로 주변 유역의 토지 피복이 산림 등으로 구성되어 있기 때문에 영양염류의 유입이 적고, 온도가 낮아 상대적으로 부영양화가 적게 일어나는 것으로 판단된다. 이러한 결과는 저수지 수질이 저수지 주변의 토지 피복 및 이용에 크게 영향을

받는다는 것을 제시해 준다. 따라서 저수지 수질과 토지 피복 및 이용 등의 상호 연관에 대한 보다 심도 깊은 연구가 요구된다.

적 요

본 연구에서는 농림부와 환경부의 전국수질측정망 자료를 이용하여 2002년 9월 전후에 조사된 전국 290개 농업용 저수지의 Chl-*a* 농도를 예측하였다. 우리나라 290개 농업용 저수지의 9월 전후 영양상태를 분류한 결과, 부영양 상태 이상을 나타내는 저수지가 TSI_{CHL} 64.1%, TSI_{TP} 75.5%로 대부분의 저수지가 높은 부영양화 상태를 보였다. 이렇게 분류된 저수지의 영양 상태를 환경특성에 따라 판별분석을 실시하였다. 그 결과 전체 판별적중률은 약 60%를 보였다. 판별분석의 결과에 정준분석을 실시한 결과, 각 그룹은 영양상태에 따라 구분이 되었으며, COD, DO, TP 등이 중요한 인자로 나타났다. 또한 MLP-PSO 모델을 이용하여 부영양화에 따른 저수지 수질을 예측한 결과 높은 예측력을 보였으며 ($r=0.831$, $p<0.05$), 민감도 분석 결과 COD와 TP가 상대적으로 가장 중요한 요인으로 작용하였으며, 고도 및 제방 높이는 음의 영향을 미치는 것으로 나타났다.

사 사

본 연구는 농림부 핵심 전략 연구과제 “생물조절(먹이 연쇄)을 통한 농업용저수지 수질관리 기법개발”(306009-03-2-CG000)에 의해 수행되었음.

인 용 문 헌

김호섭, 황순진. 2004. 알은 부영양 저수지의 육수학적 특성-계절에 따른 수질변화. *육수지* **37**(2): 180-192.
 농림부, 농업기반공사. 2000. 농업용수 수질오염원 조사 종합 보고서.
 신윤근, 전상호. 2002. 아산호의 생태학적 연구 1. 이화학적 특성과 영양상태. *육수지* **35**(3): 181-186.
 이광식, 윤경섭, 김형중, 김호일. 2003. 농업용 저수지의 부영양화와 수질관리방안. *한국환경농학회지* **22**(2): 166-171.
 이재연, 이재훈, 신경훈, 황순진, 안광국. 2007. 우리나라 농업용 저수지의 영양상태 및 수질특성. *육수지* **40**(2): 223-233.
 전지홍, 윤춘경, 함종화, 김호일, 황순진. 2002. 농업용 저수지의

물리적 인자가 수질에 미치는 영향. *육수지* **35**(1): 28-35.
 Bae, M.J., Y.S. Kwon, S.J. Hwang and Y.S. Park. 2008. Comparison of four different ordination methods for patterning water quality of agricultural reservoirs. *Korean J. Limnol.* **41**(Special issue): 1-10.
 Birge, B. 2003. PSOT: a particle swarm optimization toolbox for use with Matlab. *In: Proc. IEEE Swarm Intelligence Symposium 2003 (SIS 2003)*, Indianapolis, Indiana, USA. pp. 182-186 (url: <http://www.mathworks.com/matlabcentral/fileexchange>).
 Carlson, R.E. 1977. A trophic status index for lakes. *Limnol. Oceanogr.* **22**: 361-369.
 Chau, K.W. 2006. Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. *J. Hydrol.* **329**: 363-336.
 Clerc, M. and J. Kennedy. 2002. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation* **6**(1): 58-73.
 Goldman, J., D.A. Caron and M.R. Dennet. 1987. Nutrient cycling in a microflagellate food chain, 4. phytoplankton-microflagellate interactions. *Mar. Ecol. Prog. Ser.* **38**: 75-87.
 Harper, D. 1992. *Eutrophication of Freshwater: Principles, Problems and Restoration*. Chapman and Hall press, London.
 Havens, K.E. 2000. Using Trophic State Index (TSI) values to draw inferences regarding phytoplankton limiting factors and seston composition from routine water quality monitoring data. *Korean J. Limnol.* **33**(3): 187-196.
 Hwang, S.J., C.G. Yoon and S.K. Kwun. 2003. Water quality and limnology of Korean reservoirs. *Paddy & Environment* **1**: 43-52.
 Kennedy, J. and R.C. Eberhart. 1995. Particle swarm optimization. *In: Proc. IEEE Int. Conf. on Neural Networks.* **IV**: 1942-1948.
 Kratzer, C.R. and P.L. Brezonik. 1981. A Carlson-type trophic state index for nitrogen in Florida lakes. *Wat. Res. Bulletin* **17**: 713-715.
 Lu, R.S., S.L. Lo and J.Y. Hu. 1999. Analysis of reservoir water quality using fuzzy synthetic evaluation. *Stochastic Environ. Res. Risk Assess.* **13**(5): 327-336.
 Natarajan, U., V.M. Periasamy and R. Saravanan. 2007. Application of particle swarm optimisation in artificial neural network for the prediction of tool life. *Int. Adv. Manuf. Technol.* **31**(6): 871-876.
 Rawson, D.S. 1955. Morphometry as a dominant factor in the productivity of large lakes. *Verh. Int. Ver. Limnol.*

- 12: 164-175.
- Shi, Y. and R.C. Eberhart. 1998. Parameter Selection in Particle Swarm Optimization. *In: Pro. the 7th Int. Conf. Evol. Program.* 7: 591-600.
- Smith, V.H., E. Willen and B. Karlsson. 1987. Predicting the summer peak biomass of four species of blue-green algae (cyanophyta/cyanobacteria) in Swedish lakes. *Wat. Res. Bulletin* 23: 397-402.
- StatSoft, Inc., 2004. STATISTICA (data analysis software system), Version 7, <http://www.statsoft.com>. The Mathworks. 2001. MATLAB, Version 7.0. The Mathworks, Inc., Massachusetts.
- Wilson, D.I. 2002. Derivation of the chalk superficial deposits of the North Downs, England: an application of discriminant analysis. *Geomorphology* 42: 343-364.

(Manuscript received 31 December 2007,
Revision accepted 4 February 2008)