

다변량분석법을 활용한 농업용 저수지 수질유형분류

Classification of Agricultural Reservoirs Using Multivariate Analysis

최 은 희* 김 형 중** 박 영 석***

Choi, Eun Hee · Kim, Hyung Joong · Park, Young Suk

Abstract

In order to manage the water quality in reservoir, it is necessary to understand the temporal and spatial variation of reservoirs and to classify the reservoirs. In this research, agricultural reservoirs are classified according to physical characteristics (depth, residence time, shape of the reservoir etc) and water quality using multivariate analysis (PCA and CA).

CA (Cluster Analysis) method classify reservoirs into several groups as a similarity of the reservoirs, but it is difficult to indicate a full list to the one table.

In case of PCA (Principle Component Analysis) method, it has the advantage for the classification on the reservoirs depending on the water quality similarity and also it is useful to analyze the relationship between related factors through correlation analysis. However PCA is limited to classify into several groups based on the characteristics of the reservoirs and each user should be classified as randomly subjective according to the relative position of the reservoir in the figure.

In conclusions, compared to conventional reservoirs classification methods, both CA and PCA methods are considered to be a classification method that describes the nature of the reservoir well, but classification results has a restriction on use, so further research will be needed to complement.

I. 서론

수질오염이 심한 농업용 저수지의 수질개선

사업은 시행계획수립단계에서 현재 및 목표연도의 수질을 정확하게 예측하고 공사비를 정밀하게 산출하기 위해 복잡한 수질예측모델을 구동

* 한국농어촌공사 농어촌연구원 주임연구원 (ehchoi@ekr.or.kr)

** 한국농어촌공사 농어촌연구원 주임연구원 (iamwater@ekr.or.kr)

*** 경희대학교 생물학과 교수 (parkys@khu.ac.kr)

키워드 : PCA, CA, classification of reservoir

해야한다. 하지만 사업계획단계에서는 저수지의 개략적인 수질 변화 및 이에 따른 개략 공사비를 산출하게 되며 이를 위해서는 많은 경우 신뢰성을 가지고 있는 단순화된 수질예측모델이 필요하다. 현재 많은 종류의 저수지 수질예측모델이 개발 및 활용되고 있지만 기존 저수지 수질예측모델은 각각 서로 다른 특징을 갖고 있고, 입력변수가 다양하고 양이 많기 때문에 전문적 지식을 충분히 갖추고 있지 않은 일선 실무자들은 모델 구동에 어려움이 있고, 실무자들의 경험 및 편의에 의해 모델을 적용하기 때문에 신뢰성 저하의 우려도 있다.

이로 인해 사업계획단계별로 효율적으로 사용할 수 있는 단순화된 수질예측모델의 개발이 필요하게 되었다. 그러나 저수지가 가지고 있는 다양한 특성 및 모델별 적용시 요구되는 특수 조건 때문에 모든 저수지에 범용적으로 사용가능한 단순 수질예측모델의 개발은 어렵다. 따라서 저수지를 유형화하여 각 유형별로 적합한 수질예측모델을 개발하는 과정이 필요하다.

저수지를 유형화하기 위해서는 농업용 저수지의 시공간적 변이 특성을 정확히 이해하고, 저수지의 수질과 환경인자, 그리고 생물인자간의 상호작용 등에 대한 이해가 요구된다. 이러한 필요성에 따라 김과 황(2004)은 수심에 따른 엽록소-a(Chl-a)의 농도 특성을 비교하여 우리나라 농업용 저수지를 4개의 유형으로 나누었으며 김등(2007)은 Chl-a와 총인(T-P)에 근거하여 산출된 부영양화도지수(TSI)를 이용하여 농업용 저수지를 4개의 유형으로 구분하였다. 이들 연구는 모두 두 변수간의 상관관계에 따라 저수지 유형을 분류하는 것으로서 비교적 쉽게 할 수 있다는 장점이 있다. 그러나 실제 자연현상들을 1~2개의 변수만으로 설명하기에는 어려움이

많고 1~2개의 인자만을 이용하여 분류한 현재의 저수지의 유형분류법은 각 유형별로 저수지의 다양한 특성을 잘 나타내고 있다고 보기에는 어려움이 있다. 즉 저수지의 수질은 여러 다양한 인자들이 서로 복합적으로 작용하여 나타나는 것이므로 이러한 특성이 고려된다면 보다 유용한 저수지 유형 분류가 될 수 있을 것이다.

따라서 본 연구에서는 저수지의 유효수심, 수면적, 체류시간, 형상, 수질 등 물리·화학적 특성에 따라 저수지를 유형분류하기 위해 다변량 분석법(군집분석 및 주성분분석)을 적용하였으며, 저수지 유형분류에 대한 이들 분석법의 적용 가능성을 평가하고자 한다.

II. 재료 및 방법

1. 저수지 수질 자료

전국 호소수질측정망 자료 526개 지점 중 농업용으로 사용되는 저수지 474개 지점의 자료를 활용하여 저수지 수질에 따른 유형을 분류하였다. 측정망 자료에는 수온, pH, 용존산소(DO), 전기전도도(EC), 생물화학적산소요구량(BOD), 화학적산소요구량(COD), 부유물질(SS), 총질소(T-N), T-P, Chl-a 등이 포함되었다. 수질측정망 자료는 1999년부터 지역에 따라 연 2회-4회 조사가 되고 있으며, 본 연구에서는 저수지 수질 유형을 분류하기 위하여 이들 자료 중 2007년에 조사된 자료를 사용하였다. 각 계절별로는 봄에는 413개 저수지 자료가 활용 가능하였으며, 여름에는 337개 저수지, 가을에는 437개 저수지, 그리고 겨울에는 26개 저수지 자료를 활용하였다.

2. 자료 분석

저수지 수질은 저수지가 위치한 공간적 특성에 따라 달라질 뿐만 아니라 시간에 따라서도 변동하므로 본 연구에서는 계절별로 구분하여 분석하였다. 봄은 3~5월, 여름 6~8월, 가을 9~11월, 겨울 12~2월까지의 측정 자료로 구분하였다. 그리고 만수면적, 유역면적, 총저수량, 유효저수량, 제당높이, 제당길이, 제당표고, 준공 후 기간(나이) 등 저수지의 수문학적, 지형학적 특성 또한 기록되어 수질특성과 연관성 분석에 활용되었다. 각각의 분석 방법에 따라 자료를 봄, 여름, 가을, 겨울로 구분하여 각 계절별로 이용가능한 자료를 사용하여 분석을 실시하였으며, 분석에는 저수지 수질관리에 중요한 4인자(COD, T-N, T-P, Chl-a)를 사용하였다.

저수지의 수질특성에 따른 유형분류를 위해서 주성분분석(Principle Component Analysis, PCA), 군집분석(Cluster Analysis, CA) 등을 포함하는 다변량분석법을 이용하여 분석후 그 결과를 비교하였다.

가. 자료 전처리

환경자료는 많은 인자들 간의 복잡한 상호작용 때문에 비선형성과 복잡성을 내포하고 있다. 또한 많은 noise를 내포하고 있어 이들 자료를 분석하는데 많은 어려움이 있다. 따라서 정량적인 자료 분석에 있어서 먼저 수행되어야 할 것은 이러한 noise를 제거해 주고, 자료를 표준화하는 것이다. 이때 사용될 수 있는 것은 자료를 적절한 방법에 따라 변환하는 것이다.

$$b_{ij} = \log(x_{ij}+1) \quad (1)$$

이때 $\log(0)$ 의 문제를 피하기 위하여 변수에 1을 더하여 사용한다. 로그변환을 하는 경우 변수값이 전체적으로 매우 작은 값을 보이는 경우는 위의 방법을 따르지 않고 변형된 방법을 이용한다.

$$b_{ij} = \log(x_{ij}+d)-c, \quad (2)$$

$$c = \text{Int}(\log(\min(x)));$$

$$d = \log^{-1}(c) \text{ decimal constant}$$

이러한 로그 변환 등을 하게 되면 큰 값과 작은 값 사이의 차이를 줄여줌으로서 작은 값을 보이는 변수의 중요성을 증가시켜 줄 수 있다. 또한 대부분의 통계분석에서는 자료의 균질성 등을 필요로 하는데 이러한 조건을 충족시켜 줄 수 있다.

그리고 다변량분석에서 각인자의 중요도를 동일하게 부여하기 위하여 각 인자의 값을 최소-최대값 범위에서 다음과 같이 값을 0과 1 사이의 범위 내에서 변환한다.

$$b_{ij} = \frac{(x_{ij}-\min(x))}{\max(x)-\min(x)} \quad (3)$$

$\max(x)$; 최대값, $\min(x)$; 최소값.

나. 군집분석(CA)

CA는 주어진 인자들의 수질의 물리화학적 환경인자를 이용하여 이들의 유사성에 따라 조사지점을 몇 개의 그룹으로 나누기 위하여 사용한다. CA는 N개의 개체를 대상으로 p개의 변수들을 측정하였을 때, 이 변수들을 이용하여 N개의 개체들 사이의 유사성 또는 비유사성의 정도를 측정하여 개체들을 가까운 순서대로 연결해 나가는 분석방법이다.

분석 과정은 일반적으로 ① 변수의 측정, ② 유사성의 측정, ③ 군집화, ④ 각 군집의 성격

이나 상호관계 파악 등 4단계로 이루어진다.

CA는 방법에 따라 계층구조적 방법(hierarchical cluster)과 상호배반적 방법(disjoint cluster)으로 나눈다. 계층적 군집방법은 병합적(agglomerative) 방법과 분할적(division) 방법이 있다(Jongman et al, 1995). 병합적 방법은 가까운 개체들끼리 묶어감으로써 군집을 만들어 가는 방법이며, 분할적 방법은 반대로 먼 개체들을 나누어 가는 방법이다. 병합적인 방법은 우선 가장 가까운 2개의 개체를 묶어서 하나의 군집을 만들고, 나머지(N-2) 개의 개체는 각각 하나의 군집을 이루도록 한다. 다음은 (N-1)개의 군집 중에서 가장 가까운 2개의 군집을 묶어(N-2) 개의 군집을 만든다. 이와 같은 방법으로(N-1) 단계를 반복하면 결국 N 개의 개체가 모두 묶여서 하나의 군집을 만들게 된다(<http://www.pvalue.co.kr/>). 본 연구에서는 수질인자에 따른 각 저수지들 사이의 Euclidean 거리를 계산하고 그 유사도에 따라 Ward linkage 방법을 이용하여 그룹을 형성하는 병합적 방법을 사용하였다. 분석은 다변량분석 프로그램인 PC-ORD(4.4)를 이용하였다.

다. 주성분분석(PCA)

PCA는 대표적인 간접기울기(indirect gradient) 분석법의 하나로서 고유값(eigen value)을 계산하여 군집의 유사성에 따라 분류하는 것이다. PCA는 서로 연관이 있는 변수들($X_1, X_2, X_3, \dots, X_p$)이 관측되었을 때, 이 변수들이 가지고 있는 정보들을 최대한 확보하는 적은 수의 새로운 변수들을 생성하는 방법이다. 즉, 서로 연관이 있는 p개의 변수($X_1, X_2, X_3, \dots, X_p$)들에 대하여 주요한 관심사 중의 하나는 이 변수들이 가지고 있는 변이를 측정하는

것이다. 변이란 이 변수들이 가지고 있는 정보의 양이라고 할 수 있는데, 주성분분석에서는 원래의 변수들이 가지고 있는 변이의 양을 가장 많이 확보하는 순서대로 변수들의 선형결합을 이용하여 새로운 변수를 구하는 과정이다(<http://www.pvalue.co.kr/>). 분석은 다변량 분석 프로그램인 PC-ORD(4.4)를 이용하였다.

Ⅲ. 결과 및 고찰

1. 농업용 저수지 지역별 수질 특성

표 1은 지역별 저수지의 수리수문학·지형학적 특성을 나타낸 것이다. 울산, 대구, 인천은 각각 경남, 경북, 경기 지역으로 포함시켰다. 저수지 완공 후 기간은 대부분 40년 이상이었으며 충남, 전남 지역의 경우 표고가 낮고 제당 높이 또한 낮았다.

이화학적 수질 특성을 지역별로 보면 전체적으로 서해안, 낮은 표고를 보인 지역이 대체적으로 높은 수질오염인자 값을 보였다(표 2). 특히 충남지역은 겨울을 제외한 대부분의 기간 동안(봄-여름) 가장 높은 오염도지수를 보였다. Chl-a는 특히 여름에는 49.36mg/m³로 4계절 중 가장 높았으며 봄과 가을에는 38.56mg/m³, 34.52mg/m³로 높은 값을 보였다. 또한 T-N, T-P, COD 값 또한 높은 값을 보였다. T-P는 봄에 0.08mg/L로, COD는 여름에 10.24mg/L로 높았다. 한편 경기지역은 여름철에 수질이 크게 악화되는 양상을 보였다. Chl-a는 44.38mg/m³로 증가하였으며, COD는 10.69mg/L로 증가하여 겨울철까지 높은 값을 보였다. 경기지역에서 가을에 DO가 높은 값을 보였는데

표 1. 지역별 수리수문학적, 지형학적 특성

도	저수지 수	준공나이		만수면적		유역면적		총저수량		유효저수량		제당높이		제당길이		제당표고	
		년	표준편차	ha	표준편차	ha	표준편차	천 m ²	표준편차	천 m ³	표준편차	m	표준편차	m	표준편차	m	표준편차
강원	29	37.17	3.51	22.74	8.64	1424.69	316.12	1296.81	771.29	1258.50	626.51	21.22	1.81	240.17	95.69	184.42	16.26
경기	49	43.41	2.70	33.67	6.64	869.18	243.20	1686.65	593.36	1644.52	481.98	15.62	1.40	328.70	73.61	84.48	12.51
경남	50	41.10	2.67	19.89	6.58	885.56	240.75	1293.43	587.39	1271.51	477.14	21.07	1.38	262.70	72.87	101.79	12.38
경북	85	40.46	2.05	37.30	5.05	1963.71	184.65	3111.72	450.51	2550.00	365.95	20.71	1.06	248.23	55.89	127.73	9.50
전남	64	44.77	2.36	40.92	5.81	1352.41	212.80	2236.76	519.19	2244.55	421.73	17.80	1.22	567.95	64.41	61.61	10.94
전북	41	45.80	2.95	67.51	7.26	1393.75	265.87	3700.62	648.67	3633.18	526.91	18.15	1.53	711.85	80.47	151.86	13.67
충남	58	45.43	2.48	55.21	6.11	1254.07	223.53	2757.83	545.38	2629.03	443.01	14.92	1.28	319.86	67.66	47.95	11.50
충북	37	38.65	3.11	43.46	7.65	1710.43	279.87	3325.71	682.83	3206.40	554.66	21.81	1.61	232.38	84.71	155.33	14.39

표 2. 지역별 계절별 수질인자 차이

계절	도	저수지 수	pH		DO		EC		BOD		COD		SS		T-N		T-P		Chl-a	
			표준편차	mg/L	표준편차	µS/cm	표준편차	mg/L	표준편차	mg/L	표준편차	mg/L	표준편차	mg/L	표준편차	mg/L	표준편차	mg/L	표준편차	mg/m ³
봄	강원	31	7.98	0.11	10.07	0.24	91.84	23.43	2.00	0.21	3.85	0.47	4.13	1.02	1.41	0.19	0.01	0.01	6.34	3.99
	경기	17	7.81	0.15	7.05	0.32	107.35	31.64	2.76	0.28	5.24	0.64	5.02	1.38	1.44	0.25	0.01	0.01	7.85	5.39
	경남	46	7.90	0.09	7.55	0.19	108.13	19.23	1.55	0.17	4.45	0.39	4.59	0.84	0.98	0.15	0.01	0.01	4.59	3.28
	경북	75	8.05	0.07	10.22	0.15	129.07	15.06	2.50	0.13	5.66	0.30	5.37	0.66	0.95	0.12	0.01	0.00	8.45	2.57
	전남	61	7.46	0.08	9.54	0.17	163.79	16.70	2.15	0.15	5.20	0.34	6.21	0.73	1.94	0.13	0.03	0.00	10.58	2.85
	전북	48	7.98	0.09	9.02	0.19	172.71	18.83	2.19	0.17	6.31	0.38	6.27	0.82	2.30	0.15	0.03	0.00	19.02	3.21
	충남	52	7.49	0.09	8.37	0.18	208.32	18.09	2.95	0.16	8.86	0.37	9.10	0.79	2.87	0.14	0.08	0.00	38.56	3.08
충북	30	7.27	0.12	8.70	0.24	127.47	23.81	2.19	0.21	5.57	0.48	4.00	1.04	1.32	0.19	0.01	0.01	9.64	4.06	
여름	강원	8	8.16	0.26	10.06	0.69	112.13	54.60	3.01	0.55	6.95	1.39	12.29	4.97	0.99	0.25	0.02	0.02	14.24	10.37
	경기	39	7.49	0.12	8.41	0.31	222.59	24.73	5.58	0.25	10.69	0.63	16.72	2.25	1.40	0.11	0.05	0.01	44.38	4.70
	경남	69	8.20	0.09	8.02	0.24	126.08	18.59	2.20	0.19	5.44	0.47	9.04	1.69	0.86	0.09	0.01	0.01	8.78	3.53
	경북	45	8.13	0.11	7.77	0.29	147.68	23.02	4.09	0.23	7.63	0.59	12.41	2.10	0.81	0.11	0.02	0.01	17.48	4.37
	전남	42	8.12	0.11	7.16	0.30	203.71	23.83	3.56	0.24	7.92	0.61	12.75	2.17	1.53	0.11	0.06	0.01	29.72	4.53
	전북	54	7.69	0.10	5.57	0.27	170.76	21.02	1.91	0.21	6.99	0.54	10.17	1.91	1.81	0.10	0.06	0.01	19.81	3.99
	충남	45	7.66	0.11	7.04	0.29	250.48	23.02	3.44	0.23	10.24	0.59	13.70	2.10	2.17	0.11	0.10	0.01	49.36	4.37
충북	24	7.26	0.15	5.63	0.40	127.67	31.52	2.95	0.32	6.21	0.80	8.73	2.87	1.00	0.15	0.02	0.01	9.73	5.99	
가을	강원	29	7.83	0.11	9.65	0.29	68.40	14.33	1.47	0.24	3.84	0.45	2.61	1.50	0.84	0.16	0.01	0.01	8.95	3.34
	경기	49	6.99	0.09	10.14	0.22	157.52	11.02	4.40	0.19	6.17	0.34	7.57	1.15	1.68	0.12	0.03	0.00	26.70	2.57
	경남	50	7.85	0.09	8.44	0.22	110.41	10.91	1.91	0.19	4.57	0.34	5.76	1.14	0.86	0.12	0.02	0.00	17.47	2.54
	경북	85	7.70	0.07	8.54	0.17	99.98	8.37	2.91	0.14	6.28	0.26	6.07	0.88	0.98	0.09	0.02	0.00	16.09	1.95
	전남	64	7.75	0.08	6.34	0.19	101.47	9.65	1.96	0.16	5.56	0.30	10.68	1.01	1.88	0.11	0.05	0.00	18.76	2.25
	전북	41	7.15	0.09	6.80	0.24	123.22	12.05	1.68	0.21	6.30	0.38	10.12	1.26	2.39	0.13	0.06	0.01	19.45	2.81
	충남	58	7.80	0.08	6.88	0.20	131.37	10.13	2.76	0.17	6.54	0.32	8.54	1.06	2.36	0.11	0.06	0.00	34.52	2.36
충북	37	7.57	0.10	7.09	0.25	99.58	12.69	1.79	0.22	4.80	0.40	5.14	1.33	1.14	0.14	0.02	0.01	11.40	2.96	
겨울	경기	4	7.20	0.33	11.15	0.74	220.25	38.08	3.73	0.90	7.48	2.36	8.88	9.34	4.19	0.90	0.11	0.05	55.30	23.19
	경남	1	7.60		11.30		180.00		2.70		6.10		16.80		1.05		0.06		8.40	
	경북	1	7.50		12.60		129.00		2.20		5.10		2.50		1.02		0.02		9.30	
	전남	4	7.50	0.33	10.80	0.74	53.25	38.08	1.03	0.90	2.68	2.36	2.05	9.34	1.07	0.90	0.04	0.05	1.88	23.19
	전북	6	8.15	0.27	10.85	0.60	119.50	31.10	2.07	0.73	9.13	1.93	19.35	7.63	2.44	0.73	0.10	0.04	47.18	18.93
	충남	2	7.90	0.47	13.40	1.04	194.50	53.86	1.10	1.27	4.60	3.34	2.70	13.21	1.23	1.27	0.02	0.06	5.50	32.79

이는 높은 Chl-a로 인해 식물플랑크톤의 활발한 광합성의 결과로 보인다. 그리고 전북지역 저수지에서 다른 계절에 비해 겨울철에 높은 수질 오염 값을 보였다. 그러나 겨울철에 조사된 저수지 수가 제한되어 있기 때문에 이를 일반화 하기는 한계가 있다. 한편 충남, 전북, 전남지역에서 염소(CI)의 값이 각각 24.0, 20.1, 22.0mg/L로 5.7-8.5mg/L의 값을 보인 다른 지역에 비해 높게 나타났다. 염소의 양은 전기전도도와 매우 높은 상관관계($r=0.9$)를 보여 이들 지역에서 매우 높은 전기전도도는 염소에 의한 것으로 판단된다. 이는 이들 지역에 많은 간척지가 있고 바다와 가까이 있기 때문에 나타나는 현상으로 보인다.

2. 다변량분석

저수지 수질 관리에 있어서 가장 중요하게 다루어지는 4인자(COD, T-N, T-P, Chl-a)를 이용하여 CA 및 PCA 분석을 실시하였다.

가. CA 분석

그림 1은 계절별로 CA 분석을 실시한 결과이다. 봄과 여름에는 각각 분석에 사용된 413 및 337개 저수지는 Euclidean 거리(또는 유사도)에 따라 크게 2개 그룹, 4개의 하위그룹으로 구분되었으며, 가을과 겨울에는 각각 437, 26개 저수지가 3개의 그룹으로 구분되었다. 이와 같이 그룹을 만드는 데 대한 통계적 유의성을 검정하기 위하여 MRPP(Multi-Response Permutation Procedures)를 실시한 결과 모두 유의한 것으로 나타났다($p<0.05$). 그림에서는 많은 저수지를 제한된 공간에 나타내게 되므로 잘 표기가 되지 않는다. CA 결과의 전형적인

표현방법은 그림 1 겨울과 같이 나타내어진다. 봄 자료의 경우 그림에서 가장 오른쪽에는 신희, 전대, 기지, 대유 등 4개의 저수지가 나누어졌는데 이들은 수질오염이 가장 많이 된 곳들이다. 따라서 CA 분석결과에 따르면 저수지는 수질특성에 따라 4개의 그룹으로 나눌 수 있으며 그림에서 왼쪽에 위치한 저수지는 오염정도가 낮은 곳인 반면 오른쪽은 오염이 심한 곳을 나타내 주고 있다. 여름에는 기흥(신갈), 신희, 봉암, 덕림 저수지 등이 오염된 저수지로 같이 분류되었으며, 가을에는 부남호, 덕림, 왕송, 신류, 초대, 덕동 등이 동일 오염그룹으로 분류되었다. 한편 겨울에는 다른 계절에 높은 오염도를 보였던 덕림, 기지, 기흥과 함께 여천 저수지도 같은 그룹으로 분류되었다. 전체적으로 신희, 기지, 덕림, 기흥 등의 저수지가 가장 오염된 곳으로 구분된다.

이 CA 방법은 저수지를 유사도에 따라 몇 개의 그룹으로 구분해 줄 수 있어 유용하게 사용될 수 있으나 전체 목록을 하나의 표로 나타내는데 있어 어려움이 있다. 또한 각 그룹에 대한 인자들의 값이 어떠한 특성을 가지는가를 나타내는 데는 한계점이 있다.

나. PCA 분석

저수지 수질 자료를 사계절로 나누고 저수지 수질을 PCA 분석을 실시한 결과 전체적으로 계절별로 유사한 양상을 보였다(그림 2). 봄의 경우 가로축과 세로축은 각각 73.4%(Eigen value=2.936) 및 15.2%(Eigen value=0.608)의 전체 분산 보였다. 그리고 여름의 경우는 가로축과 세로축이 각각 68.1%(Eigen value =2.724) 및 20.9%(Eigen value= 0.837), 겨울의 경우는 가로축이 62.1%(Eigen value

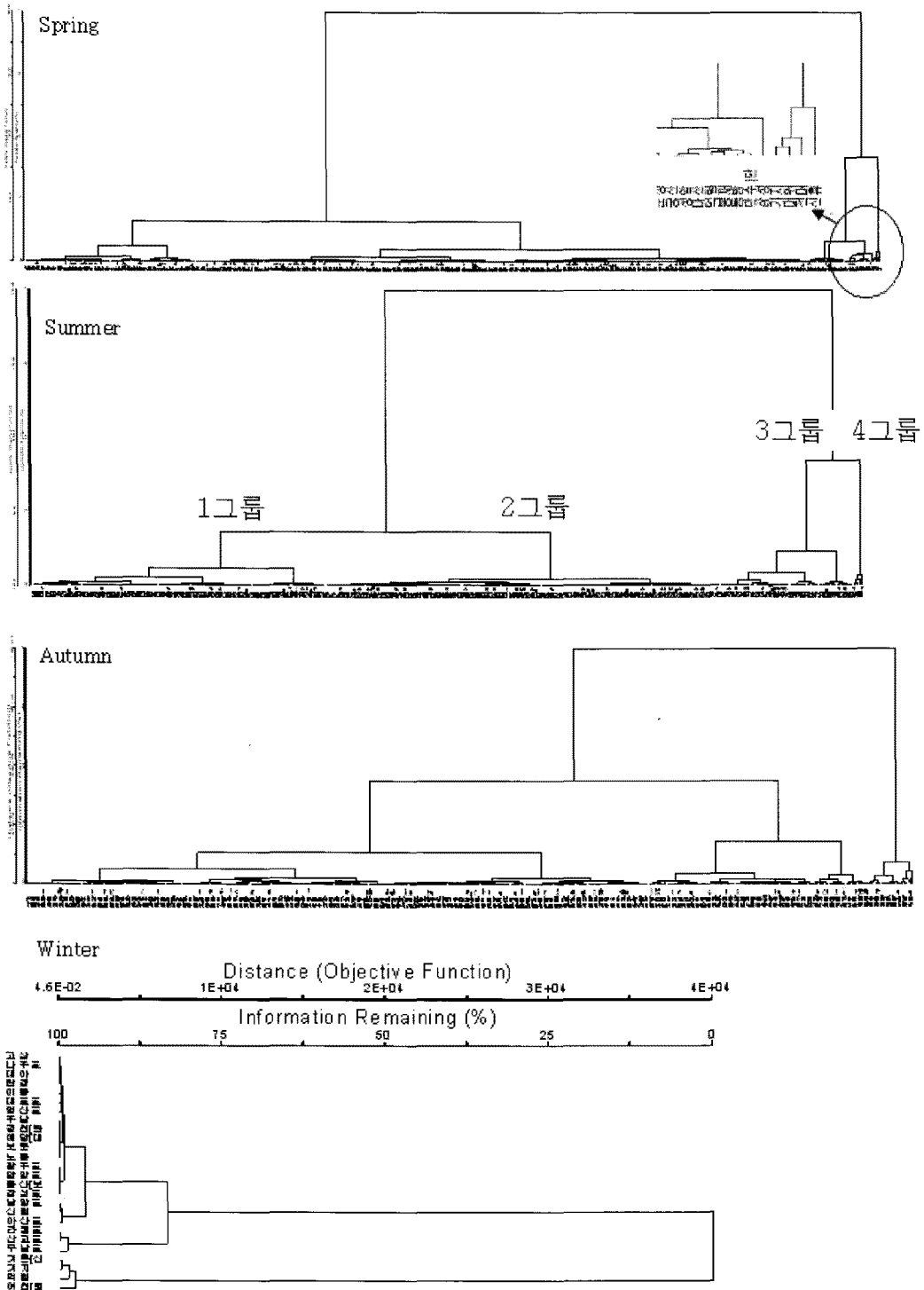


그림 1. CA분석에 따른 수질인자 특성별 저수지 분류

=2.483), 세로축이 20.8%(Eigen value =0.832), 겨울은 가로축이 75.5%(Eigen value= 3.018), 세로축이 16.1%(Eigen value=0.646)의 분산도를 보여주며 자료를 잘 구분하여 주었다.

4계절 모두 수질오염이 많이 된 저수지들은

왼쪽에 위치하며 오른쪽에는 상대적으로 오염이 적은 저수지들이 위치하고 있다. CA분석에서 수질오염이 많이 된 저수지들이 하나의 그룹으로 구분된 것과 같이 이들은 PCA 상에서 왼쪽에 많이 흩어져 나타났다. 봄의 경우 신희, 전대, 기지, 대유 등이 왼쪽에 위치하는데 이들은

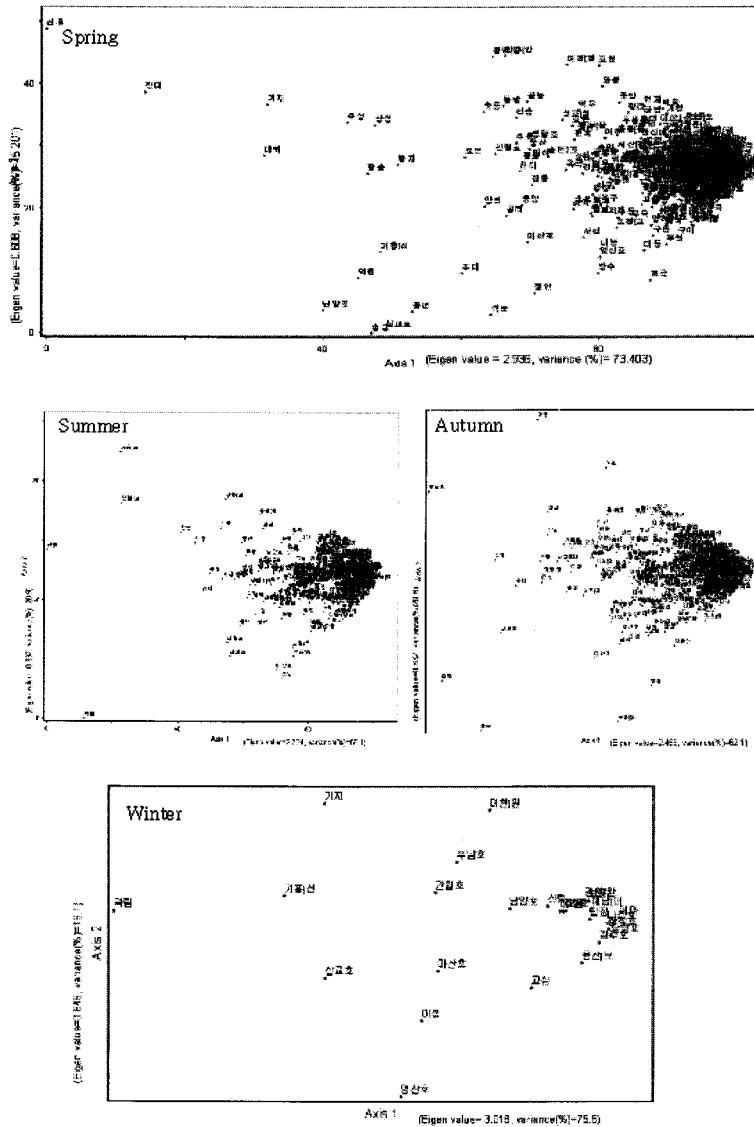


그림 2. PCA를 이용한 계절별 저수지 수질 특성분석

가장 왼쪽에 위치한다. 신희저수지는 Chl-a가 237.0mg/m³, T-P 0.26mg/L, T-N 6.2mg/L, COD 26.0mg/L 등으로 높았으며, 기저저수지 또한 Chl-a 167.4mg/m³, T-P 0.98mg/L, T-N 4.8mg/L, COD 16.5mg/L로 높은 값을 보였다. 한편 여름에는 기흥(신갈), 신희, 봉암, 덕림 저수지들이 가장 오염된 것으로 분류되었으며, 가을에는 부남호, 덕림, 왕송, 신희, 초대, 삼교호, 덕동 저수지 등이 오염이 심하며, 겨울에는 덕림, 기저, 기흥 저수지, 삼교호가 높은 오염도를 보이는 것으로 분류되었다. 신희는 겨울을 제외한 다른 계절에 모두 높은 오염도를 보이는 그룹으로 분류되었으며, 덕림은 여름, 가을, 겨울에 오염도가 심하였다. 전반적으로 다소 변이는 있지만 오염도가 심한 저수지는 연중 매우 높은 오염도를 보인다는 것을 보여준다. 한편 이들 오염도가 가장 심한 저수지들은 지리적으로 기흥(신갈) 저수지 등과 같이 도시발달 및 인구밀집과 높은 관계가 있다는 것을 보여준다.

PCA의 각 축에 대한 수질인자의 상태적인 설명력을 평가하기 위하여 각 저수지의 각 축에 대한 좌표값과 그에 대한 수질인자 값 사이의 상관계수를 계산하였다(표 3). 그 결과 가로축은 T-P가 -0.885로 가장 높은 상관성을 보였으며 세로축은 T-N이 -0.572로 가장 높은 상관성을 보였다. 이는 PCA 축에서 오른쪽에

위치한 저수지 일수록 T-P의 값이 낮고, 반대로 왼쪽에 위치한 저수지일수록 높은 값을 보이며, 위쪽에 위치한 저수지 일수록 T-N의 값이 낮고, 아래쪽에 위치한 저수지 일수록 높다는 것을 의미한다. 따라서 저수지는 수질 특성에 따라 오염된 곳과 그렇지 못한 곳을 잘 구분해 주고 있다. 여름 및 가을의 경우도 각 축에 대한 상관계수에는 차이가 있으나 봄과 유사한 특성을 보였다. 한편 겨울철에는 가로축이 COD와 가장 높은 상관성($r=-0.909$)을 보였으며 세로축은 Chl-a와 높은 상관성($r=0.516$)을 보였다. 세로축은 위쪽에 있을수록 Chl-a값이 높다는 것을 의미한다.

PCA 상에서 이러한 수질인자의 분포특성은 그림 3과 같이 나타내어 질 수 있다. 그림 3은 여름 자료의 PCA 결과에서 가로축에 대하여 COD가 높은 상관성을 가지고 세로축에 대해서는 T-N이 높은 상관성을 가진다는 것을 보여주고 있다. 그림에서 각각 저수지를 나타내는 원의 크기는 각 인자 값의 상대적인 값의 차이를 나타내는 것으로서 원이 클수록 높은 값을, 원이 작을수록 낮은 값을 나타낸다. 기흥저수지는 가로축에 대하여 COD 값이 매우 높지만 T-N값은 낮은 값을 보이는 반면, 덕림 저수지는 가로축에 대하여 COD 값은 낮지만 T-N 값은 가장 높은 값을 보이는 것을 보여주고 있다.

이러한 특성은 저수지의 수질을 나타내는데

표 3. PCA 축에 대한 각 수질인자의 상관계수

변수	봄		여름		가을		겨울	
	가로축	세로축	가로축	세로축	가로축	세로축	가로축	세로축
COD	-0.88	0.361	-0.855	0.412	-0.797	0.410	-0.909	0.276
T-N	-0.778	-0.572	-0.663	-0.706	-0.684	-0.670	-0.857	-0.424
T-P	-0.885	-0.191	-0.886	-0.215	-0.870	-0.219	-0.886	-0.352
Chl-a	-0.337	0.113	-0.877	0.350	-0.789	0.408	-0.821	0.516

그러나 CA 분석 결과를 PCA와 결합하여 나타낸다면 그룹을 결정하는데 있어서 다소 도움을 받을 수는 있을 것으로 판단된다.

IV. 결론

농업용 저수지 수질개선사업 사업계획단계별 적용가능한 단순화된 수질예측모델을 개발하기 위해 유사한 특성을 가진 저수지들의 유형화작업을 수행하였다. 현재 개발되어 있는 저수지의 유형분류법은 COD, Chl-a, 수심 등 1~2개의 인자간의 상관관계에 따라 저수지 유형을 분류하고 있어 비교적 쉽게 적용할 수 있는 장점이 있는 반면 실제 자연현상들을 1~2개의 변수만을 설명하기에는 어려움이 있고 1~2개의 인자만을 이용하여 분류한 현재의 저수지의 유형분류법은 각 유형별로 저수지의 다양한 특성을 잘 나타내고 있다고 보기에는 어려움이 있다. 그래서 저수지의 유효수심, 수면적, 체류시간, 형상, 수질 등 저수지가 가지고 있는 다양한 변수들을 고려한 다변량분석법을 적용하여 저수지 유형분류를 수행하였다.

다변량분석법 중 주성분분석(PCA) 및 군집분석(CA)법을 활용하여 농업용 저수지의 유형을 분류하였으며, 그 결과 CA 방법은 저수지를 유사도에 따라 몇 개의 그룹으로 구분해 줄 수 있어 유용하게 사용될 수 있으나 전체 목록을 하나의 표로 나타내는데 있어 어려움이 있으며, 또한 각 저수지 그룹에 대한 인자들의 값이 어떠한 특성을 가지는가를 나타내는 데는 한계점이 있었다.

PCA분석법을 통한 농업용 저수지 유형분류는 각 저수지의 수질 유사성에 따라 저수지를

잘 구분해 주는 장점이 있으며 또한 상관관계분석 등을 통하여 관련 인자들 간의 연관성을 분석하는데 유용하게 사용될 있지만 CA 분석에서와 같이 분석대상 저수지를 특성에 따라 몇 개의 그룹으로 나누는 것은 제한적이며, 이를 위해서는 사용자가 그림 상에서 각 저수지의 상대적인 위치를 보고 임의로 주관적으로 구분하여야 하는 문제점이 있음을 확인하였다.

따라서 기존의 1~2개의 인자를 활용한 저수지 유형분류법에 비해 CA 및 PAC 분석법 모두 저수지가 가지고 있는 다양한 특성을 대변하는 유형분류법으로의 유용성은 확인하였으나, 유형분류 결과의 활용도(도표화 및 그룹화)에 제한성을 가지고 있는 것으로 판단되며, 이를 보완할 수 있는 연구가 추후 필요할 것으로 판단된다.

이 연구는 농어촌연구원 자체자금 연구비 지원에 의해 수행되었습니다.

참고문헌

1. 김호섭, 황순진. 2004. 육수학적 특성에 따른 국내 저수지의 부영양화 유형분석-엽록소 a와 수심을 중심으로. 한국육수학회지 37(2): pp.213~226.
2. 김호섭, 최은미, 김동우, 공동수, 김경만, 김범철. 2007. 농업용 저수지 유역환경특성에 따른 수질경향 분석. 한국하천호수학회지 40(2): pp.214~222.
3. 농업기반공사. 2005. 효율적인 농업용 저수지 용수 확보 및 이용방안 연구. 농어촌연구원.
4. 전지홍, 윤춘경, 함중화, 김호일, 황순진. 2002. 농업용 저수지의 물리적 인자가 수질에 미치는 영향. 육수지 35(1): pp.28~35.
5. Jongman, R. H. G., Braak, C. J. F. Ter, Van Tongeren, O. F. R. Data Analysis in Community and Landscape Ecology, Cambridge University Press., 1995.
6. <http://www.pvalue.co.kr/>