

저수지 수질조사 지점간 유사성 분석

A Study on Measuring the Similarity Among Sampling Sites in Lake

이요상*, 고덕구**, 이현석***

Yosang Lee, Deuk-koo KOH, Hyun Seok Lee

Abstract

Multivariate statistical approaches to classify sampling sites with measuring their similarity by water quality data. For empirical study, data of two years at the 9 sampling sites with the combination of 2 depth levels and 7 important variables related to water quality is collected in reservoir. The similarity among sampling sites is measured with Euclidean distances of water quality related variables and they are classified by hierarchical clustering method. The clustered sites are discussed with principal component variables in the view of the geographical characteristics of them and reducing the number of measuring sites. Nine sampling sites are clustered as follows; One cluster of 5, 6, and 7 sampling sites shows the characteristic of low water depth and main stream of water. The sites of 2 and 4 are clustered into the same group by characteristics of hydraulics which come from that of main stream. But their changing pattern of water quality looks like different since the site of 2 is near to dam. The sampling sites of 3, 8, and 9 are individually positioned due to the different tributary.

Key words : water quality, similarity, clustering method, principal component analysis

1. 서론

수질측정망은 전국 하천 및 호수 등 수질보전대상 공공 수역에 대한 수질현황을 종합적으로 파악하여 수질변화 추세를 파악하고 이미 집행된 주요 정책 사업의 효과를 분석 하여 장래 수질보전정책수립을 위한 기초 데이터를 확보하기 위해서 운영된다. 국내에서는 1970년대 후반부터 수질측정망 단위의 정기적 수질측정이 이루어지기 시작하였으며 2008년에는 1,476개 지점으로 운영되고 있다. 지금까지는 새로운 수질측정망 선정이나 기존 수질측정망 변경 요인이 발생하면 수질 관리자의 경험과 주관적 판단에 의해 이루어지고 있었다. 이로 인하여 수질이 유사한 지역이 중복되거나 다른 측정지점과는 수질 패턴이 상이한 지점이 누락되는 경우가 발생하게 되는 한계와 문제점을 내포하고 있다. 또한 수질데이터는 충분한 자료의 취득이 상대적으로 어렵고 대상항목이 다양하며 수질항목에 따른 영향이 상이하게 나타나기 때문에 수질측정망에 관련된 연구는 활발하지 않은 실정이다.

본 연구에서는 댐저수지의 9개 수질관측지점, 각 2개의 수심에서 2년 동안 관측된 7개 수질항목 데이터를 이용하여 관측지점간 수질 패턴 유사성을 유클리디안 거리로 측정하고 이를 바탕으로 계층적 군집분석에 의해 유사 관측지점을 분류하였다. 분류된 관측지점의 특성을 주성분 분석과 수질 관점에서 평가하여 과학적이며 효율적인 수질측정망지점선정에 활용 가능한 방법을 제시하고자 한다.

* 정회원 · 한국수자원공사 K-water연구원 수자원연구소 책임연구원 · E-mail : yslee@kwater.or.kr

** 정회원 · 한국수자원공사 K-water연구원 수자원연구소 소장 · E-mail : dkkoh@kwater.or.kr

*** 정회원 · 한국수자원공사 K-water연구원 수자원연구소 공동연구원 · E-mail : leehs2005@kwater.or.kr

2. 기본이론

유사성에 의한 개체분류는 관심변수들에 의해 개체 간 유사성을 측정하고 이를 이용하여 유사한 개체를 분류하는 방법으로 군집분석(clustering analysis)이라 한다. 군집분석은 주어진 데이터의 구조에 의해 사전지식 없이 개체에 대한 의미 있는 정보를 찾아내는 탐색적 자료분석으로 유사성이 높은 개체 순으로 개체를 묶어가는 계층적(hierarchical) 방법과 일정수준의 유사성을 기준으로 개체를 묶거나 사전에 결정된 군집의 개수에 의해 개체를 분류하는 비계층적 방법이 있다. 계층적 군집분석 방법은 사용이 용이하나 개체가 일단 어느 군집에 소속되면 다른 군집으로 이동될 수 없고 이상 개체는 제거되지 않고 반드시 어느 군집에 속하게 된다. 비계층적 군집분석은 일정한 군집 수준을 보장하나 사전에 군집 수를 결정해야 사용이 가능하다. 그러므로 군집분석의 일반적인 순서는 계층적 군집분석 방법에 의해 적절한 군집 개수를 결정하고 비계층적 방법에 의해 개체를 분류하게 된다. 군집변수의 수가 분류하려는 개체 수에 비해 상대적으로 큰 경우 비계층적 방법은 적절하지 못하다. 주성분분석(principal component analysis : PCA)은 선형상관 관계가 있는 다수의 변수들을 서로 독립인 주성분변수로 변환하여 데이터의 차원을 줄이는 다변량분석방법이다. 주성분변수는 원 변수들의 선형결합으로 계산된다. 원변수들의 공분산행렬로부터 얻은 고유치에 대응하는 고유벡터(eigen vector)를 구하여 이를 선형계수로 사용한다. 주성분변수는 서로 독립이다. 누적 설명기여율이 80%인 주성분변수만을 활용하여 2차분석을 함으로써 원 데이터의 차원을 축약하게 되며, 주성분변수 이름은 선형계수의 크기에 의해 부여되어 선형계수를 부하값이라고도 한다.

3. 연구지역

본 연구의 대상인 용담댐은 현재 4지점의 수질조사지점이 운영되고 있으나, 조사지점간 수질 유사성 분석을 위해서는 수질관측지점을 늘려 저수지 전역에 관측 위치가 고밀도로 분포되도록 관측지점을 선정하였다. 저수지 내의 수질측정 지점은 그림 1 과 같이 댐 앞에 1번 지점에서부터 시작하여 주자천 유입 지류에 2번, 3번 지점을 설정하였고, 저수지 본류에 5개 지점(4번~7번, 9번)을 선정하였으며 진안천 지류에 1개(8번) 지점을 선정하였다(Fig. 1). 각 지점별 수질변화 관측을 위해 2005년과 2006년에 걸쳐 1월부터 12월까지 월 1~4회씩 각 관측지점에서 19개 수질항목에 대한 수질 조사를 실시하였다.

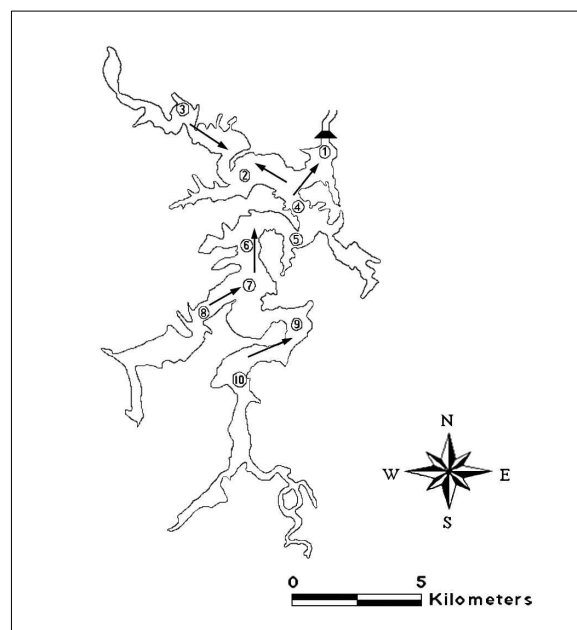


Fig. 1. Sampling sites at Yongdam reservoir

4. 적용 및 분석

4.1. 저수지 수질변화

투명도의 변화를 살펴보면, 저수지에서 주흐름이 이루어지는 1, 2, 4, 5, 6, 7지점은 봄철에 투명도가 가장 좋게 나타나고 하절기에 투명도가 감소하는 전형적인 저수지 수질변화패턴을 유지하나, 지류유입 하천에 영향을 많이 받는 3, 8, 9지점은 년중 투명도가 크게 개선되지는 않았다.

저수지에서 SS는 평수기에 전지점에 걸쳐 큰 변화를 보이지는 않으나, 강우기에 유입되는 탁수가 장기간 저수지에 머무는 동안에는 높은 SS농도를 나타냈다. 특히 금강분류가 유입되는 주 흐름인 9번지점에서 가장 높게 나타나며 하류로 가면서 7번지점과 6번지점에서는 점차 농도가 낮아지다가 4번지점 부근에서는 거의 침강 소멸되는 경향을 나타냈다.

COD_{Mn}와 T-N의 경우에는 전형적인 댐저수지의 계절별 수질변화를 나타냈으며, T-P와 PO₄-P도 거의 유사한 계절별 수질변화를 나타냈으나 하절기에 SS에 부착되어 유입되는 특성으로 인해 SS처럼 하절기에 높게 측정되었고 하류로 갈수록 급격히 농도가 낮아지는 현상을 나타내고 있다. Chl-a도 전형적인 댐저수지의 계절별 변화양상을 나타내었다.

4.2 관측지점 유사성 분석

본 연구에서의 실증분석은 수질을 군집변수로 하여 관측위치에 대한 군집분석을 실시하였다. 관측위치간 유사성은 주요 수질변수들에 의해 측정된 후 각 개체들은 평균연결방법에 의해 계층적 군집화하고, 이 과정에서 얻어진 덴드로그램에 의해 적절한 군집개수를 결정하였으며 군집개수 설정의 적절성 및 분류된 군집의 특성 파악을 위하여 상관계수행렬을 사용하는 주성분분석을 적용하였다.

Fig. 2 는 13개 수질변수를 군집변수로 하여 관측지점 9개의 유사성을 표준화 유클리디안 거리로 계산하고 평균연결법에 의해 얻은 덴드로그램이다. 덴드로그램의 수평 점선 (2005년 유사성 수준=74, 2006년 유사성 수준=72)은 주관적인 판단에 의해 설정한 군집분류 기준선이다. Table 1 은 덴드로그램에서 설정된 기준선의 유사성 수준과 관측위치 군집 결과를 정리한 것이다. □, △, ▽, ○로 표시한 군집기호는 특성이 동일한 지점을 표시한 것이며, 지점 5, 6, 7은 측정된 수질 면에서 2년 모두 유사한 지점으로 판단되었고 지점 3은 2005년에는 지점 2와 2006년에는 지점 4와 유사한 것으로 분류되었다.

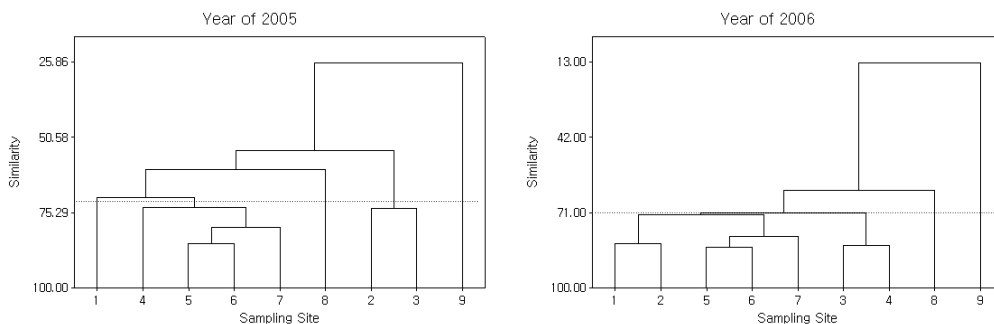


Fig. 2. Dendrogram by Average Linkage

Table 1. Result of Clustering Analysis

| Year | Similarity level | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------|------------------|---|---|---|---|---|---|---|---|---|
| 2005 | 74 | | △ | △ | | ○ | ○ | ○ | | |
| 2006 | 72 | □ | □ | ▽ | ▽ | ○ | ○ | ○ | | |

관측지점 2는 2005년에는 관측지점 3과 2006년에는 관측지점 1과 유사한 지점으로 분석되었다. 한편 관측지점 8과 9는 2년 모두 상이한 관측지점으로 분류되었다. 그러므로 군집분석 결과 저수지 특성을 보이는 지점 (5, 6, 7), 하천적 특성을 갖는 지점 (1, 2, 3, 4), 그리고 상류 지류가 다른 지점 두 지점 8과 9을 하나의 권역으로 간주하여, 즉 4개 권역에서 적절한 실험설계 방법에 의해 관측 지점과 관측 빈도를 결정하면 된다.

Table 2 는 군집분석 결과 얻어진 군집의 특성 파악을 위하여 13개 특성 변수의 상관관계수 행렬을 이용하여 실시한 주성분분석 결과이며 Fig. 13 은 묶여진 군집(Table 1 의 군집 기호 사용)을 그룹 표시로 하여 제 1 주성분과 제2 주성분 산점도를 그린 것이다. 주성분변수 2의 크기는 부하의 크기가 큰 SD, COD_{Mn}상, Chl-a 하(음의 부호), Chl-a상(음의 부호, 2005년만)에 의해 주로 결정된다. 음의 의미는 Chl-a 값은 SD나 COD_{Mn}는 음의 상관관계를 가짐을 의미한다. 주성분 1은 주성분 2의 크기에 영향을 미치지 않은 나머지 수질 변수들에 의해 크기가 결정된다.

Table 2. Result of principal component analysis(PCA)

| Parameter | Year of 2005 | | Year of 2006 | |
|---------------------------|------------------------|-----------|--------------|-----------|
| | First PC ^{*)} | Second PC | First PC | Second PC |
| SD | -0.26 | 0.35 | -0.27 | 0.40 |
| SS upper | 0.31 | 0.15 | 0.31 | 0.02 |
| COD _{Mn} upper | 0.12 | 0.54 | 0.10 | 0.79 |
| T-P upper | 0.31 | 0.14 | 0.31 | 0.05 |
| PO ₄ -P upper | 0.29 | 0.28 | 0.30 | -0.10 |
| T-N upper | 0.29 | 0.19 | 0.26 | 0.16 |
| Chl-a upper | 0.07 | -0.44 | 0.30 | 0.11 |
| SS lower | 0.32 | 0.01 | 0.30 | 0.06 |
| COD _{Mn} lower | 0.30 | 0.08 | 0.27 | 0.03 |
| T-P lower | 0.32 | -0.01 | 0.31 | 0.07 |
| PO ₄ -P lower | 0.31 | -0.08 | 0.29 | 0.08 |
| T-N lower | 0.31 | -0.21 | 0.27 | -0.03 |
| Chl-a lower | 0.25 | -0.43 | 0.26 | -0.38 |
| Eigen value | 9.28 | 2.09 | 10.29 | 1.42 |
| Cumulative percentage (%) | 71.4 | 87.5 | 79.1 | 90.0 |

*) PC : 주성분(Principal Component)

2005년 주성분 산점도에서 관측지점 9번은 제1 주성분 수질에 의해 다른 관측지점들과 전혀 다른 패턴을 보이고 있으며 제2 주성분 수질에 의하면 3번과 8번지점이 완전히 독립적으로 위치하고 1, 2, 4, 5, 6, 7번 지점이 비슷한 거동을 보이고 있다. 이중에서 1, 4, 5, 6, 7번 지점은 분류 흐름구간이어서 수질특성 면에서 유사한 속성을 보이는 것이 가능하며 2번 지점은 용담댐 도수터널의 방류량이 많아 4번까지 유입된 흐름이 2번 지점까지 영향을 미치기 때문에 분류의 수질특성을 나타내기 때문으로 판단된다. 2006년의 9번 지점은 2005년과 같이 제1 주성분 수질에 의해 완전히 독립적으로 거동한다. 제2 주성분 수질에 의하면 8번 지점은 다른 지점들과 완전히 독립적으로 거동하며 3번과 2번지점이 독립적으로 거동한다. 결론적으로 1,4,5,6,7번 지점이 수질특성 면에서 유사하다는 것을 의미한다.

5. 결 론

용담댐 9개 지점에서 7개 수질항목 특성에 대하여 상·하층 2개 관측위치에서 측정된 데이터를 활용하여 월별 수질변화 특성을 정리하였으며, 이를 바탕으로 관측지점간 유사성을 분석하였다.

수심이 깊고 저수지의 주 흐름을 형성하는 1, 4, 5, 6, 7번 지점은 저수지적 수질 특성을 나타내는 것으로 평가되었다. 특히 5, 6, 7번 지점은 측정값이 유사한 것으로 나타났으나, 1번과 4번은 수질변화 경향은 유사하나 수질값은 상이한 것으로 나타났다. 2번 지점은 본류흐름에 영향을 받으나 수문특성에 따라 큰 영향을 받을 경우에는 저수지 본류의 수질특성을 나타내고, 본류의 영향을 적게 받는 해에는 본류와 상이한 수질특성을 나타내는 지점으로 평가되었다. 3, 8, 9번 지점은 유입하천에 영향을 직접 받아, 본류수질과 상이한 특징을 나타낼 뿐만 아니라 각각 유입하천의 오염도가 다르기 때문에 서로 다른 수질특성을 나타내는 것으로 분석되었다. 9번 지점은 제1 주성분에 의해 다른 관측지점과 상이하였고, 관측지점 3, 8은 제2주성분(SD, COD_{Mn}상, Chl-a 수질 특성) 특성에 의해 다른 관측지점과 상이함을 보였다.

이상과 같이 다변량 통계기법을 활용한 측정 지점별 분석을 실시하였으며, 각 지점별 수질특성에 따른 유사성 분석결과를 제시하였다. 분석된 결과는 활용목적에 따라 새롭게 측정망을 구축하는데 적용할수 있으며, 이를 통해 수집된 자료는 저수지 수질현황 파악을 정확하고 용이하게 하고 장래 수질보전정책 수립을 위한 기초자료로 활용이 가능할 뿐만 아니라 다양한 연구나 업무에 활용 될 수 있을 것으로 판단된다.

참 고 문 헌

- 건설교통부, 2004. 수문관측 매뉴얼
- 한국수자원공사, 2005, 2006, 2007. 용담댐 일원 하천유량측정 등 수문조사보고서
- Johnson, Dallas E., 1998, Applies Multivariate Methods for Data Analysis, Duxbury Press.
- Johnson, Richard A. and Wichern Dean W., 2007, Applies Multivariate Statistical Analysis, 6th edition, Pearson Prentice Hall.
- Tanaka, Y., Tarumi, T., and Huh, M. H., 1994, Research and Applications of Quantification Methods, in New Approaches in classification and Data Analysis, Springer-Verlag, Berlin.