

다변량 통계분석을 이용한 낙동강 상수원수의 수질변화 특성 조사

김경아[†] · 김예진* · 송미정 · 지기원 · 유평종 · 김창원*

부산광역시 상수도사업본부 수질연구소

*부산대학교 환경공학과

Evaluation of Water Quality Characteristics in the Nakdong River using Multivariate Analysis

Gyungah Kim[†] · Yejin Kim* · Mijeong Song · Keewon Ji · Pyungjong Yu · Changwon Kim*

Water Quality Institute, Busan Water Authority

*Department of Environmental Engineering, Pusan National University

(Received 7 September 2007, Accepted 31 October 2007)

Abstract

This study was estimated water quality to raw water quality management of the Maeri intake station in the Nakdong River using Multivariate Analysis. The results of Principle Component Analysis was explained up to 76.9% of total water quality by three principle components. The 1st, 2nd was explained 44.7%, 17.9% and third was explained 14.3%. Also, the three factors was derived from Factor Analysis. The 1st factor was estimated as the metabolism and organic matter pattern related to algal growth. The 2nd factor was judged as the pollution of pattern related to the discharge from stream of the Nakdong River and 3rd factor was viewed as the hydrological variation pattern related to particle matter. The results of Cluster Analysis were classified into three groups.

keywords : Cluster analysis, Factor analysis, Multivariate analysis, Principle component analysis

1. 서론

낙동강 유역은 크게 4개의 중권역(상류권역, 중류권역, 하류서부권역, 하류동부권역)으로 구분되는데, 상류권역에는 강원 태백시 일부와 경북 구미시 등 16개 시·군의 전부 혹은 일부, 중류권역에는 대구광역시와 경산시 등 20개 시·군·구의 전부 혹은 일부, 하류서부권역에는 전북 남원시 일부와 경남 진주시 등 13개 시·군의 전부 혹은 일부, 하류동부권역에는 부산광역시와 경남 밀양시 등 16개 시·군·구가 포함되어 있다(부산광역시, 2002). 낙동강 유역의 각 권역은 농촌지역, 도시지역, 산업공단지역 등 다양한 형태의 오염원 배출환경이 조성되어 있으며 이들 오염물질들은 낙동강으로 유입되어 수질을 악화시키고 있다. 또한 강우시에는 심각한 오염발생의 가능성이 더 높아지므로, 비강우시의 수질 특성과 계절적인 특성이 조합되어 낙동강의 상수원수의 수질은 매우 다양한 패턴을 보일 뿐 아니라 주기적으로 심각한 부영양화로 인해 정수처리에 어려움을 겪고 있다.

그러므로 상수원수의 수질관리 및 상수원수의 수질에 대한 정수장의 운전변수 변경 등 정수장의 운영 효율을 최적화하기 위해 지속적인 수질 모니터링을 통한 정확한 수

질평가가 요구되어지며 수질평가 결과는 정수장의 운전변수의 조정에 있어 가치있는 기준이 될 수 있을 것으로 사료된다. 따라서 본 연구에서는 다변량분석법을 이용하여 복잡한 수질형태를 가지고 있는 매리원수의 수질특성을 해석하고 평가하였으며, 수질형태를 분류하기 위해 군집분석을 사용하여 매리원수의 수질특성을 3가지로 분류하였다.

다변량분석법은 일반적으로 연구대상으로부터 측정된 2개 이상의 변수들의 관계를 동시에 분석할 수 있는 모든 통계적인 기법을 말하는 것으로 주성분분석, 인자분석, 회귀분석, 다변량분산분석, 판별분석, 군집분석 등 여러 가지 기법이 있으며(김, 1998) 이러한 통계적 기법들은 수질의 종합평가에 널리 사용되고 있다. Ohta 등(1986)은 일본의 三河灣에서 주성분분석을 이용하여 해역의 수질특성을 파악하고 적조 발생인자를 도출하는 연구를 하였고 박 등(2005)은 상관분석과 요인분석으로 팔당호에서의 수역별 계절별 조류관련 인자를 선정하고 요인별 점수로 조류발생 정도를 예측할 수 있는 선형회귀모형을 유도하는 연구를 하였는데 전체적으로 조류발생에 대한 모형의 설명력은 높지 않은 것으로 나타났다. 또한 박 등(1986)은 다변량분석에 의한 금호강의 수질을 평가하였고 오 등(1984)은 낙동강의 23개지점에서의 수질자료를 활용하여 주성분분석으로 하천수질을 분석하였는데 단기간에 걸친 비교적 적은 수질 측정값을 이용했기 때문에 신뢰성 있는 해석을 할 수 없어

[†] To whom correspondence should be addressed.
rana@busan.go.kr

보다 장기간에 걸친 자료의 획득이 요구되어진다 하겠다. 이외에도 다양한 연구논문들이 발표되었다(권 등, 1974; 박 등, 2001; 신 등, 1998; 장, 1999).

이에 본 연구에서는 낙동강하류에 위치해 있으며 항상 수질오염가능성과 다양한 수질패턴을 보이는 부산시 상수원수인 매리원수를 장기간에 걸친 수질자료와 다양한 통계분석을 통해 좀 더 정확하게 수질을 평가해 보고자 하였다. 따라서 2001~2005년의 최근 5년간 매리원수의 수질현황을 파악하고 원수수질과 댐방류량 및 강수량과의 상관관계를 조사 평가하였으며, 주성분분석(Principle Component Analysis)과 인자분석(Factor Analysis)을 적용하여 원수 수질항목의 종합특성치(주성분)의 추출과 수질항목간 상관관계를 조사하였다. 그리고 수질항목간에 동일한 요인에 의해 영향을 받는 항목들을 파악해보고, 항목간의 군집화 및 인자와의 상관관계를 시도해 봄으로써 수질변화 특성을 보다 쉽게 해석하고 평가해 보고자 하였다. 또한 군집분석(Cluster analysis)를 이용하여 수질변화 특성이 몇 가지로 유형화되어 발생하는지를 알아보고 향후 수질변화에 능동적으로 대처하기 위한 기초자료로 활용하고자 하였다.

2. 재료 및 방법

2.1. 연구자료

매리원수의 수질특성을 파악하기 위하여 2001년부터 2005년까지의 매주 수질 자료의 월별 평균 통계자료를 이용하였다. 수질조사항목으로 pH, DO, BOD, COD, 탁도, SS, NH₃-N, NO₃-N, T-N, T-P, 클로로필-a이며 수질오염공정시험방법(환경부, 2004)으로 분석한 자료들이다. 그리고 본 연구논문에 이용된 강수량은 기상청(2006) 홈페이지에 게재된 자료를 인용하였으며, 댐방류량은 국가수자원관리종합정보시스템(2006)의 홈페이지에 게재된 자료를 활용하였다.

2.2. 자료의 통계적 처리

2.2.1. 주성분분석(Principle Component Analysis)

주성분분석이란 다변량해석의 여러 가지 방법 중에서도 기본적인 방법으로서 상호간에 상관성을 갖는 여러 종류의 특성치들의 정보를 큰 손실없이 서로 무상관한 소수개의 종합특성치(주성분)로 요약해 내는 분석기법이다(오 등, 1984). 주성분분석법(혹은 고유벡터방법)의 가장 큰 특징은 다량의 자료의 요약표현이 가능하다는 점 이외에도 여러 개의 상호 유관한 특성치들의 변화양상으로부터 이들 속에 내재하고 있는 상호독립적인 고유의 패턴(즉 공유벡터)을 도출할 수 있다는 점이다(백 등, 1982).

즉, 다변량으로 구성되는 데이터로부터 소수의 특징적인 변량을 합성하여 주성분을 산출하고, 차원이 축소되어 데이터 분석이 용이해진 주성분을 사용하여 전체 다변량 데이터의 경향을 분석하게 된다.

2.2.2. 인자분석(Factor Analysis)

많은 변수가 서로 복잡하게 상관하고 있는 경우에도 그

들의 상관관계를 설명할 수 있는 몇 가지의 공통적인 성분을 생각하여, 개개의 변수를 그러한 공통적인 성분을 반영하는 부분과 각각의 변수에 독자적인 부분으로 나누어서 생각할 수 있다. 인자분석에서는 공통적인 성분을 공통인자(common factor) 또는 단지 인자(factor)라 부르고, 독자적인 성분을 독자인자(unique factor)라고 부른다. 인자분석의 목적은 변수간의 상관행렬로부터 공통인자를 끄집어내어 그 공통인자를 이용해서 변수간의 상관관계를 설명하고, 공통인자와의 관계에 의해서 각 변수의 성질을 간결한 형태로 기술하는 것이다.

인자분석의 목적이 특정변수간의 상호관련을 통해서 잠재적인 구조를 탐색하는데 있는 것이라면 이 목적을 위해서 인자부하행렬이나 인자구조행렬의 추정치 등으로부터 인자의 해석이 실시된다. 변수들이 여러 인자에 대하여 비슷한 인자부하량을 나타낼 경우에 변수들이 어느 인자에 속하는지를 분류하기가 어렵다. 따라서 변수들의 인자부하량이 어느 한 인자에 높게 나타나도록 하기 위해서 인자축을 회전시킨다. 회전방식은 몇 가지가 있는데 크게 직각회전(orthogonal rotation)과 사각회전(oblique rotation)으로 나뉘어진다. 직각회전방식에는 Varimax, Quartimax, Equimax 등이 있는데 이 중에서 Varimax방식이 가장 널리 이용된다. 본 연구에서는 Varimax 직각회전방식을 사용하였다.

2.2.3. 군집분석(Cluster Analysis)

군집분석의 목적은 주어진 많은 수의 관측개체를 몇 개의 그룹(군집)으로 나눔으로써(즉 세분화함으로써) 대상집단을 이해하고 군집을 효율적으로 활용하고자 하는 것이다.

군집분석은 기법에 따라 크게 다음과 같이 두 가지로 분류된다. 하나는 계층적 방법(hierarchical method)이다. 이 방법에서는 군집의 형성에 위계가 있어서 일단 한 군집에 속하게 된 두 개체는 다시 흩어지지 않는다. 다른 하나는 비계층적 방법(non-hierarchical method)이다. 이 방법에서는 군집이 형성된 이후에도 일정 기준에 따라 개체들이 이합집산을 되풀이하게 된다. 한편 군집분석은 관측개체간 거리, 즉 비유사성을 어떻게 정의하느냐에 따라 또 달라진다.

계층적 군집방법은 거리가 가까운 개체부터 차례대로 묶어 그룹을 형성해 나가는 응집된 계층방법(Agglomerative Hierarchical Method : AHM)과 전체대상을 하나의 군집으로 묶은 후 개체들을 차례로 크기가 더 작은 여러 개의 군집들로 나누는 분할된 계층방법(Devisive Hierarchical Method : DHM)이 있다. 계층적 군집방법은 군집간 거리가 어떻게 정의되느냐에 따라 ① Nearest-Neighbor Method(최단연결법) ② Furthest-Neighbor Method(최장연결법) ③ Centroid Method(중심연결법) ④ Between-groups linkage(군집간 평균연결법) ⑤ Within-groups linkage(군집내 평균연결법) ⑥ Ward's Method(와드의 방법) 등이 있다. 비계층적 군집방법은 일반적으로 사용되는 계층적 군집분석과 달리 군집화 과정이 순차적으로 이루어지지 않는 군집분석법을 일컫는다. 비계층적 군집화 방법을 실행하기 위해서는 중심을 기준으로 군집의 수와 최초의 시작점을 지정해야 한다. 비계

층적 군집화 방법을 일반적으로 K-평균 군집분석법이라고 하며, 본 연구에서는 비계층적 군집화 방법에서 K-평균 군집분석법이 가장 많이 사용되고 있으므로 이를 사용하여 데이터를 군집화하였다. K-평균 군집분석법은 군집화의 각 단계가 끝나면서 발생하는 오류를 계산하여 오류가 발생하지 않는 방향으로 군집화를 계속해 나가는 것이 특징이다.

3. 결과 및 고찰

3.1. 연도별 수리수문학적 특성 및 원수 수질변화 현황 조사

Fig. 1의 낙동강 유역도를 살펴보면 주요 유입 지천 중 남강이 하천연장 및 유역면적에 있어서 가장 크고, 그 다음이 금호강, 반변천의 순이며 낙동강 본류 상류부에는 안동댐, 지류인 남강에 남강댐, 황강에 합천댐, 반변천에 임하댐이 위치해 있다. 따라서 매리원수 수질에 영향을 미칠 수 있는 수리수문학적 인자인 유량과 관련된 이들 4개댐의 최근 5년간(2001~2005년)의 총방류량에 대해 조사하고 인근 4개지역의 강수량 현황을 병행 조사하여 수질에 미치는 영향을 조사해 보았다. 또한 매리원수의 수질항목별 변화양상도 추가로 살펴보았다. 2001년도에는 가장 적은 강수량으로 인해 댐방류량 또한 가장 낮게 나타났으며 강수량이 최고값을 보인 2003년도에는 역시 댐방류량도 최고였다. 2002년과 2003년도에는 집중호우로 인하여 강수량이 높게 나타났는데 2002년 8월에 총강수량이 2829.6 mm로 최고치를 보였으며 그 다음으로 2003년 7월에 2109.6 mm를 나타내었다.

또한 5년간의 6~9월 기간 중에 내린 강수량이 연강수량의 68.1%를 차지하여 하계에 집중되어 있음을 알 수 있었다.

최근 5년간의 수질변화를 살펴보면 강수량이 적은 해인

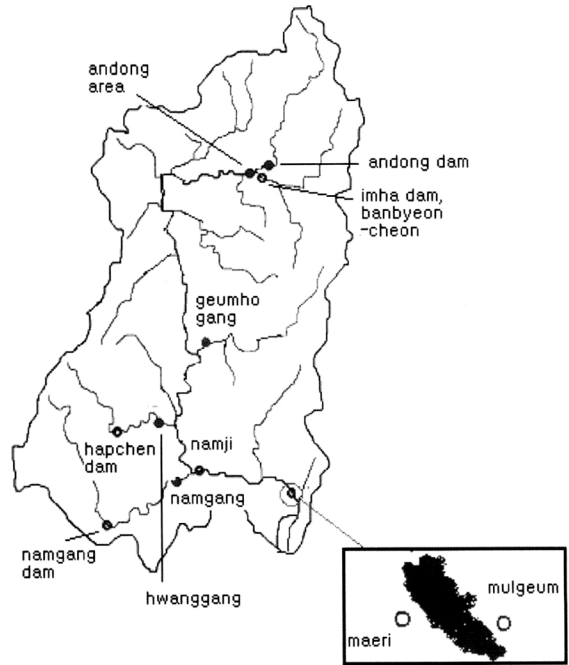


Fig. 1. Drainage basin of the Nankdong river.

2001년은 주요지점별 낮은 강수량으로 인해(Table 1 참조) 탁도와 SS가 가장 낮은 반면 유기물과 클로로필-a의 양이 상대적으로 높은 것을 알 수 있다. 반면에 2003년은 집중호우로 인하여 5년간 대비 최대강수량을 보였으며 따라서 탁도와 SS가 최고치였으며 유기물과 클로로필-a의 양은 상대적으로 최저치를 나타내었다. 반면 탁도의 경우 2002년 연평균값이 27.9 NTU인데 반해 8월과 9월의 평균 탁도가 각각 62.8 NTU와 137.6 NTU를, 2003년도는 연평균값이 35.0 NTU이며 8월과 9월의 평균 탁도가 각각 35.8 NTU와 134.7 NTU를 보였다. 이것은 최고 강수량 후 댐수량 관리

Table 1. Average data of rainfall and total discharge for each area (rainfall : mm, discharge : m³/sec)

Year	Total discharge	Total rainfall	Rainfall			
			Pusan area	Andong area	Hapcheon area	Jinju area
2001	78.1	362.1	97.6	68.2	94.9	101.4
2002	142.2	573.0	173.8	107.2	140.6	151.4
2003	197.9	660.9	194.0	131.7	155.2	180.0
2004	156.2	481.2	115.5	108.8	123.1	133.8
2005	112.1	375.1	115.1	73.5	93.3	93.3

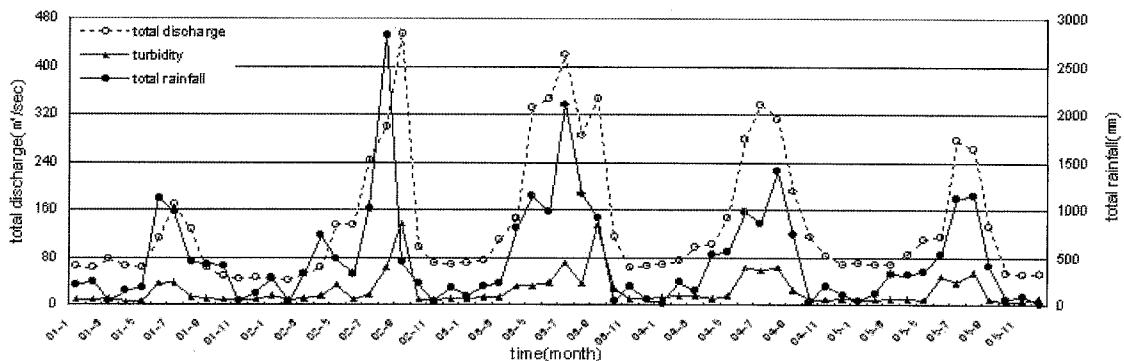


Fig. 2. Variation of turbidity of raw water of Maeri by total discharge and total rainfall.

차원에서 댐방류량을 증가시킴으로써 초기방류량의 증가로 인해 탁도가 오히려 높아지는 경향을 보인 것으로 사료된다.

Table 3은 강수량과 방류량의 변화에 따른 수질과의 연관성을 알아보기 위한 상관도 분석 결과이다. 상관계수를 살펴보면 총강수량을 기준으로 볼 때 주요 4개 지점별 상관계수는 진주지역이 0.989로 타 지역보다 높게 나타났으며, 이는 남강권이 전체 낙동강의 유량 증가에 기여하는 부분이 다른 지역에 비해 크다는 사실을 입증한다. 총 강수량 대비 pH의 상관계수는 -0.606으로 음의 상관관계를, 탁도와는 양의 상관관계를 보였다(R=0.577). 총방류량 대비 각 수질항목간의 상관관계에서는 pH가 -0.730의 높은 음의 상관관계를, BOD와도 역시 -0.631로 비교적 강한 음의 상관관계를 보였으나 탁도와는 0.832의 높은 양의 상관관계를 보였는데 이것은 7~8월의 집중호우가 원수의 고탁도를 유발함을 입증한다. Fig. 2에서도 보이는 바와 같이, 총강수량과 총방류량, 그리고 원수의 탁도의 변화 경향이 같음을 확인할 수 있다. Table 3은 총강수량이 각 수질인자와 가지는 상관도보다 총방류량이 각 수질인자와 가지는 상관도가 항상 더 크다는 것을 보여주고 있어, 원수의 수질에 영향을 주는 인자는 강수량 보다는 상류댐의 총방류량이라는

사실을 알 수 있다. 이것은 앞서 언급한 최고강수량 후 댐 방류량의 증가로 인한 탁도 증가와도 관련이 있는 것으로 여겨진다.

3.2. 주성분분석을 이용한 주성분 추출 및 수질항목 간 상관관계 비교

최근 5년간의 수질자료를 토대로 서로 상관되어 있는 수질변수들이 가지고 있는 복잡한 현상을 데이터축소를 통한 주성분 추출방법을 이용하여 요약하는 주성분분석을 우선 행하였다. 이 주성분분석은 서로 상관되어 있는 다변량들이 가지고 있는 주요성분의 추출을 목적으로 하며 저차원그래프를 이용하여 개체들을 분류한다.

우선 Table 4에서 보는 바와 같이 수질항목간의 상관관계를 조사하였다. BOD는 클로로필-a와 상관계수가 0.860으로 매우 높은 양의 상관성을 보였으며 또한 pH와도 0.818로 매우 높은 양의 상관성을 보였다. 이것은 조류발생으로 유기물 증가에 따른 결과로 보이며 또한 조류발생시 조류가 광합성작용을 통하여 이산화탄소를 이용하므로 조류발생시 pH값이 높아지는 결과 때문으로 추정된다. SS와 탁도는 상관계수가 0.835로 매우 상관성이 높은 것으로 확인되

Table 2. Variation of characteristics of raw water quality (Maeri intake plant, from 2001 to 2005)

		Turbidity	Temperature	pH	BOD	COD	SS	DO	NH ₃ -N	NO ₃ -N	Chloro-phylla	T-N	T-P
2001	Min	5.1	2.5	7.7	2.1	5.3	10.0	7.1	0.017	1.3	23.8	2.338	0.034
	Max	37.5	29.0	9.2	4.0	6.9	21.6	16.0	0.518	3.7	138.3	4.827	0.102
	Average	13.2	16.3	8.6	3.1	6.1	16.2	10.9	0.152	2.3	66.5	3.206	0.062
2002	Min	8.8	2.7	7.7	1.4	4.0	10.7	7.1	0.026	1.8	8.5	1.911	0.007
	Max	137.6	25.4	9.2	4.8	7.4	108.6	14.7	0.393	3.6	95.2	5.880	0.152
	Average	27.9	15.2	8.3	2.6	5.5	25.6	10.6	0.126	2.5	44.7	3.371	0.092
2003	Min	9.7	1.8	7.1	1.3	4.6	12.3	7.2	0.024	1.9	7.7	2.122	0.025
	Max	134.7	25.5	8.7	3.5	6.2	172.5	13.5	0.498	3.2	78.5	4.182	0.163
	Average	35.0	14.8	7.9	2.1	5.2	36.4	10.1	0.122	2.5	39.7	3.138	0.068
2004	Min	8.0	3.2	7.4	1.1	3.6	14.2	6.9	0.000	1.9	15.5	2.133	0.021
	Max	62.5	27.9	9.3	4.4	7.4	121.1	14.3	0.142	3.0	165.0	3.837	0.096
	Average	25.0	16.5	8.4	2.6	5.0	33.0	10.4	0.052	2.5	59.3	2.820	0.061
2005	Min	5.3	3.0	7.6	1.8	3.7	12.2	7.2	0.006	1.3	18.0	1.577	0.028
	Max	54.2	26.4	9.4	4.4	6.9	94.5	14.6	0.195	3.5	107.4	3.349	0.112
	Average	18.2	15.7	8.5	2.7	5.3	25.7	10.4	0.050	2.2	44.1	2.544	0.073

Table 3. Correlation comparison of rainfall and discharge and water quality

Correlation	Total rainfall	Total discharge	Rainfall				pH	Turbidity	BOD	COD	NH ₃ -N	DO	NO ₃ -N	Chloro-phylla	T-P	T-N
			A	H	J	P										
Total rainfall	1.000	.753	.965	.972	.989	.936	-.606	.577	-.448	-.013	-.149	-.661	-.403	-.462	-.357	-.034
Total discharge		1.000	.744	.755	.753	.659	-.730	.832	-.631	-.166	-.207	-.662	-.352	-.560	-.392	-.060
Rainfall	A		1.000	.939	.949	.855	-.622	.598	-.468	-.053	-.148	-.666	-.408	-.456	-.371	-.020
	H			1.000	.968	.847	-.576	.620	-.457	.001	-.138	-.628	-.382	-.439	-.369	-.040
	J				1.000	.903	-.603	.597	-.439	.004	-.149	-.648	-.391	-.446	-.359	-.028
	P					1.000	-.548	.422	-.376	-.013	-.142	-.616	-.381	-.448	-.287	-.039

A : Andong area; H : Hapcheon area; J : Jinju area; P : Pusan area

Table 4. Correlation comparison of items of water quality

Correlation	pH	BOD	COD	SS	DO	Turbidity	NH ₃ -N	NO ₃ -N	Chlorophyll-a	T-P	T-N
pH	1.000	.818	.455	-.418	.651	-.608	-.036	.198	.771	.219	-.090
BOD		1.000	.731	-.342	.661	-.506	.167	.256	.860	.349	.051
COD			1.000	.022	.358	.020	.067	.021	.654	.211	.026
SS				1.000	-.353	.835	-.112	-.194	-.302	-.269	.128
DO					1.000	-.510	.409	.730	.753	.654	-.022
Turbidity						1.000	-.152	-.337	-.443	-.350	.078
NH ₃ -N							1.000	.492	.142	.633	.114
NO ₃ -N								1.000	.405	.739	.063
Chlorophyll-a									1.000	.422	-.125
T-P										1.000	.146
T-N											1.000

Table 5. Total variance explanation by principal component analysis

Component	Initial eigenvalues			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	4.918	44.706	44.706	4.918	44.706	44.706
2	1.964	17.851	62.557	1.964	17.851	62.557
3	1.576	14.329	76.886	1.576	14.329	76.886
4	.955	8.681	85.567			
5	.621	5.647	91.214			
6	.355	3.225	94.439			
7	.205	1.862	96.301			
8	.148	1.342	97.643			
9	.124	1.128	98.771			
10	.082	.741	99.512			
11	.054	.488	100.000			

었다. SS는 불용성현탁입자의 양을 나타내며 탁도가 높을 수록 SS도 높아지므로 이러한 결과가 산출된 것으로 판단 된다. 그리고 pH는 SS, 탁도와 각각 음의 상관성(R=-0.418, -0.608)을 보였으며 탁도와 BOD도 서로 음의 상관관계(-0.506)로 나타났다. 이러한 결과는 강우시 고탁도 발생으로 SS와 탁도가 증가하는 반면 pH는 낮아지며 또한 탁질의 영향과 강우로 인한 오염물의 희석으로 생물학적으로 분해가능한 유기물의 감소에 의한 결과로 보인다. 그리고 SPSS(Statistical Package the Social Science)를 이용하여 주요 11개 수질항목에 대한 주성분분석을 실시하고 그 결과를 Table 5에 나타내었다. 주성분의 수를 결정하는 데는 여러 가지 방법이 있으나, 본 연구에서는 고유값(Eigenvalue)이 1.0이상인 값을 갖는 주성분 축만을 고려하였다. 고유값은 도출된 주성분이 설명하는 분산의 크기를 나타내는 것으로서 고유값이 1.0보다 크다는 것은 하나의 주성분이 변수 1개 이상의 분산을 설명한다는 것이다. 따라서 고유값이 1.0이상인 3개의 인자를 추출하였다. Table 5에서 전체 11개 수질항목간의 상관계수를 통해서 구한 각 축의 고유값과 기여율을 살펴보면 제1인자의 기여율은 44.7%로 제1인자만으로도 전체수질의 44.7% 정도를 설명하고 있으며, 제2인자를 포함하면 62.6%까지 해석할 수 있을 것이다. 따라서 3개의 인자를 추출하여 전체 수질의 약 76.9%를 설명할 수 있을 것으로 사료된다.

3.3. 인자분석을 이용한 수질특성 평가

각 수질인자들의 특성을 파악하기 위하여 주성분추출방법을 사용하여 인자를 추출한 후 인자별 점수를 얻는 인자분석을 실시하였다. 인자와 수질변수(항목)와의 관계를 보다 정확하게 설명하기 위하여 Varimax법(직교회전방법)으로 회전시켜 인자분석을 실시하였고 Table 6과 Fig. 3에 결과를 나타내었다. 직교회전된 결과(Rotated Component Matrix)에서 인자부하량을 0.70를 기준으로 살펴보면 추출된 3개의 주성분 중 제 1 주성분은 pH, BOD, COD, 클로로필-a가 높은 부하량을 가지고 있어 식물플랑크톤의 성장에 따

Table 6. Rotated component matrix by Factor analysis

	Component		
	1	2	3
pH	.787	.020	-.471
BOD	.914	.169	-.228
COD	.870	-.003	.268
SS	-.105	-.150	.884
DO	.589	.632	-.301
Turbidity	-.239	-.246	.885
NH ₃ -N	-.014	.807	.065
NO ₃ -N	.134	.852	-.155
Chlorophyll-a	.878	.244	-.232
T-P	.219	.878	-.099
T-N	-.017	.259	.338

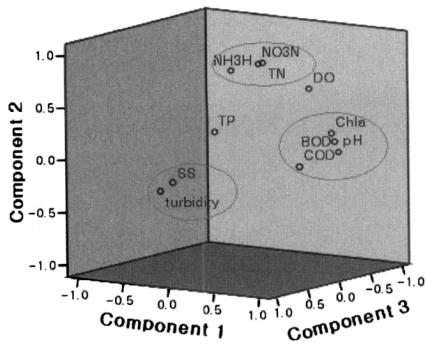


Fig. 3. Component plot in rotated space.

른 광합성 관련 물질대사 및 유기물 관련 인자가 영향을 미치고 있다고 해석할 수 있으며 제 2 주성분은 암모니아 성질소, 질산성질소, 총인이 높은 인자부하량을 보여 낙동강의 지천을 통해 유입되는 영양염류와 관련된 오염물질인자의 영향을 나타내는 주성분으로 정의할 수 있겠다. 그리고 제 3 주성분은 SS, 탁도가 높은 인자부하량을 나타내어 입자성물질과 관련된 수리수문학적인자의 영향을 나타내는 주성분으로 설명할 수 있을 것이다. 따라서 매리원수의 수질은 물질대사 및 유기물 인자가 44.7%로 가장 큰 영향을 미치는 것으로 나타났으며 오염물질인자는 17.9%, 수리수문학적인자가 14.3%로 나타났는데 박 등(2001)이 낙동강 주요지점에서의 인자부하량을 조사한 것과 비슷한 결과를 나타내었다.

Fig. 4에 의하면, 2001년도는 제1인자의 영향력이 가장 높은 것으로 나타났는데 낮은 강수량으로 인한 하천유량의 감소로 인하여 유기물 부하량이 증가했기 때문인 것으로 추측된다. 2002년도와 2003년도는 잦은 강우로 인한 강수량 증가로 낙동강의 지천을 통해 유입되는 영양염류와 관련된 오염물질인자와 수리수문학적인자의 특징을 보이는

제2인자와 제3인자의 혼합지배형을 보이는 것으로 판단된다. 2002년도에는 제2인자가 조금 더 우세하였으며 2003년도에는 제3인자의 우세를 확인할 수 있었는데 2002년보다 2003년도의 연강수량과 평균강수량이 조금 더 높았기 때문으로 보인다. 2004년도와 2005년도는 전반적으로 요인부하량이 적은 값을 보였으며 요인별 영향력도 뚜렷한 차이를 찾을 수 없었다. 이것은 아마도 적당한 강수량과 안정적인 유량의 유지로 수질변화가 크게 일어나지 않았기 때문으로 추측된다. 월별 각 인자들의 특성 조사에서는 2004년도를 제외하고 대체적으로 1~3월까지는 제2인자의 영향력이 큰 것으로 나타났는데 이것은 봄철 적당한 강수량으로 인한 낙동강 지천의 영양염류 유입 결과로 추측되며 2004년도에는 봄철 낮은 강수량으로 오히려 제1인자의 부하량이 더 높게 나타난 것으로 사료된다. 11월과 12월은 제1인자의 부하량이 다른 인자들보다 높은 것으로 나타났는데 이는 대체적으로 연중 가장 낮은 강수량과 방류량으로 인한 하천유량의 감소로 인해 나타난 결과인 것으로 판단된다. 따라서 신 등(1998)에 의한 낙동강 하구 해역의 수질평가에서 COD, 클로로필-a와 관련있는 제1인자의 변동이 크며 이는 강수량 변동에 따른 하천의 유량변동과 관련이 있다는 보고와 비슷한 결과를 나타내었다.

3.4. 군집분석을 이용한 수질유형 도출

주성분 분석과 인자분석에 사용되었던 월별 데이터의 수질항목들을 대상으로 군집분석을 수행하여 본 결과 Fig. 5와 같은 결과를 얻을 수 있었다. K-평균 군집분석방법은 대표적인 비계층적 군집분석방법으로써, 초기 군집의 수를 지정하는 단계에서부터 군집화가 시작되게 된다. 본 연구에서는 강우와 비강우시의 데이터 패턴이 매우 상이한 점으로부터 2001년에서 2005년까지의 수질유형을 3가지로 분류

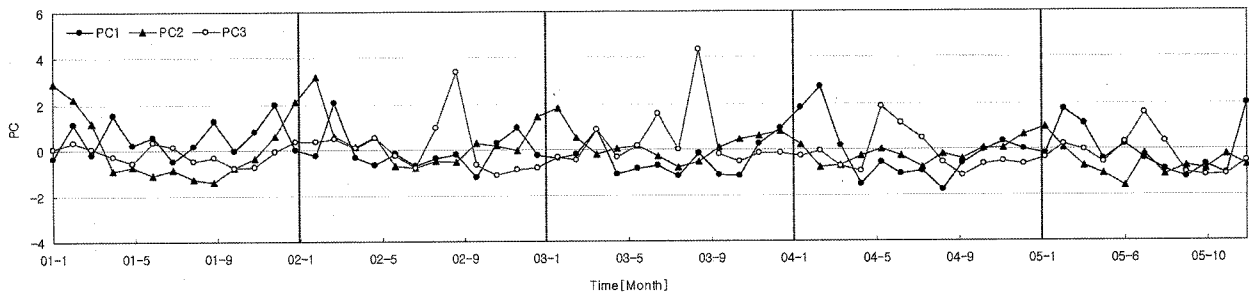


Fig. 4. Water quality variation patterns according to principle component changes.

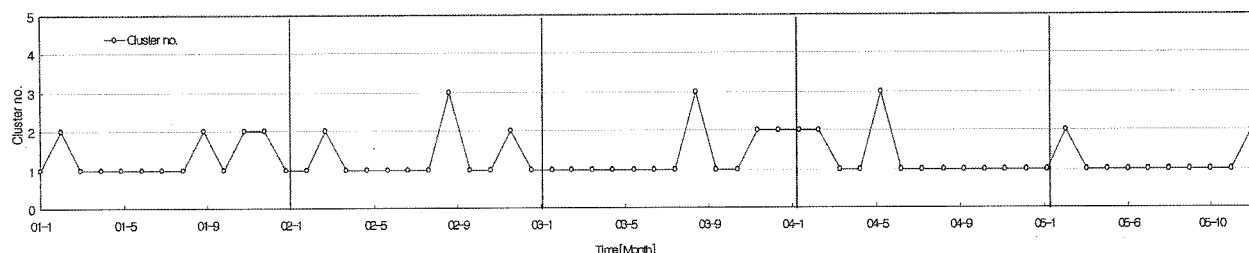


Fig. 5. Classification of water quality variation patterns by cluster analysis.

하는 것이 가장 적당한 것으로 판단되어, 군집의 수를 3가지로 지정하여 수행하였다. Fig. 4와 Fig. 5에서 보는 바와 같이, Fig. 4에서 제 3 주성분 유형에 해당하는 월과 Fig. 5의 제 3 군집에 해당하는 월(2002년 9월, 2003년 9월, 2004년 6월)이 일치하는 것을 볼 수 있었다. 이것은 2002년과 2003년의 7,8월 집중강우로 인한 댐방류량의 증가로 생긴 고탁도 때문으로 여겨진다. 또한 2004년은 2003년의 겨울철과 2004년 봄철에 걸친 장기간의 갈수기로 인해 수질이 악화되어 수질향상 차원에서 댐방류량을 증가시킴으로써 고탁도를 유발했기 때문으로 추측되어진다. 따라서 수리수문학적 인자에 의한 탁도의 증가를 대표하는 제 3 주성분의 값이 상승할 때마다 해당 월의 데이터가 제 3 군집에 속하는 것을 알 수 있었다. 그리고 pH, BOD, COD, 클로로필-a의 상승을 대표하는 데이터를 제 2 군집으로 분류하였는데, Fig. 4의 월별 제 1 주성분의 인자부하량이 1이상인 월과 Fig. 5의 제 2 군집에 해당하는 월이 대체로 비슷하게 일치하는 결과를 나타내었다. 그러므로 제 1 주성분은 조류발생에 따른 유기물 증가와 관련되며, 이것은 봄철과 겨울철 갈수기 때, 특히 3월과 12월은 조류증가와 일치하는 결과를 보여 제 2 군집은 조류발생으로 인한 수질패턴으로 해석할 수 있을 것으로 사료된다. 그 외 나머지 구간, 즉 강우가 수질을 교란하지 않는 일상적인 데이터들을 제 1 군집으로 분류했다. 결론적으로 수질 유형을 3가지로 군집화하여 보았을 때 주성분분석에 의한 인자들의 변동이 잘 반영된 결과를 도출할 수 있었으며 매리원수의 수질이 하천유량에 크게 영향을 받고 있는 것으로 판단되어졌다.

4. 결론

본 연구에서는 낙동강 하류 매리원수의 수질관리를 위해서 다변량분석법을 이용하여 수리학적측면에서의 영향성 조사와 수질항목들간의 상관관계 및 군집화를 시도해 봄으로써 수질특성을 해석하고 평가하였으며 다음과 같은 결론을 얻을 수 있었다.

- 1) 총강수량 대비 주요 4개지점별 강수량과의 상관계수는 진주지역이 0.989로 양의 상관관계를 보이며 타 지역 보다 높게 나타났다. 또한 총강수량과 pH와의 상관계수는 -0.606으로 음의 상관관계를, 탁도와는 양의 상관관계를($R=0.577$) 보였다.
- 2) 총방류량 대비 각 수질항목간의 상관관계에서는 pH가 -0.730로 높은 음의 상관관계를, BOD와도 역시 -0.631로 비교적 좋은 음의 상관관계를 보였다. 반면 탁도와는 0.832의 높은 양의 상관관계를 보였다.
- 3) 대체적으로 총강수량 보다는 총방류량이 수질항목과의 상관관계가 더 높은 것으로 나타나 강수량보다는 댐방류량이 수질에 더 영향을 끼치는 것으로 판단되었다.
- 4) 주성분분석으로는 3개의 인자를 추출하여 전체 수질의 약 76.9%를 설명할 수 있을 것으로 판단되었다. 수질항목간의 상관관계에서는 BOD는 클로로필-a 및 pH와는 각각 상관계수가 0.860, 0.818로 매우 높은 양의 상관성

을 보였다. SS와 탁도의 상관계수도 0.835로 서로 매우 상관성이 높은 것으로 확인되었다. 반면 pH는 SS, 탁도와 각각 음의 상관성($R=-0.418, -0.608$)을 보였으며 탁도와 BOD도 서로 음의 상관관계(-0.506)로 나타났다.

- 5) 인자분석으로는 전체 수질특성 중 제1인자인 물질대사 및 유기물 인자가 44.7%, 제2인자인 오염물질인자는 17.9%, 수리수문학적인자가 14.3%인 것으로 확인되었다. 또한 BOD와 제1인자간에는 0.914, 암모니아성질소는 제2인자와 0.807, 탁도와 제3인자는 0.885로 모두 매우 높은 양의 상관관계를 보였다.
- 6) 연도별 각 인자들의 특성을 살펴본 결과 2001년도는 제1인자의 영향력이 가장 높은 것으로 나타났다. 2002년도와 2003년도에는 제2인자와 제3인자가 혼합되어 지배하는 것을 알 수 있었으며 2004년도와 2005년도에는 전반적으로 인자별 영향력이 비슷하게 나타났다.
- 7) 월별 각 인자들의 특성 조사에서는 2004년도를 제외하고 대체적으로 1월~3월까지의 제2인자의 영향력이 컸으며, 2004년도에는 봄철 매우 낮은 강수량으로 인해 오히려 제1인자의 영향력이 더 높게 나타났다. 그리고 11월과 12월은 연중 가장 낮은 강수량과 방류량으로 인한 하천유량의 감소로 인해 제1인자의 부하량이 높은 것으로 판단된다.
- 8) 월별 데이터에 관한 군집분석을 수행한 결과 주성분분석 및 인자분석의 결과와 비슷한 패턴을 가지는 3개의 군집유형을 생성시킬 수 있었다. 군집분석의 결과, 강우 및 댐방류로 인한 탁도가 유발된 경우의 수질패턴과 광합성 영향인 조류발생이 활성화된 경우의 수질패턴, 그리고 그렇지 않은 일상적인 경우의 수질패턴이 군집화될 수 있었다.

참고문헌

- 국가수자원관리종합정보시스템, <http://www.wamis.go.kr> (accessed May. 2006).
- 권속표, 정용, 유진수, 신민용, 김정태, 한강하류 수질에 대한 인자분석 및 중회귀 분석, *한강상수도협회지*, 2, pp. 6-17 (1974).
- 기상청, <http://www.kma.go.kr> (accessed May. 2006).
- 김종구, 다변량해석기법에 의한 금강하구역의 수질평가, *한국환경과학회지*, 7(5), pp. 591-598 (1998).
- 박영규, 이철희, 정휘수, 다변량분석에 의한 금호강의 수질평가, *영남대학교 환경문제연구소보*, 6(1), pp. 101-117 (1986).
- 박해식, 박경길, 송미경, 백경훈, 신성교, 통계분석법을 이용한 낙동강 수질 특성 평가, *한국물환경학회지*, 17(6), pp. 693-701 (2001).
- 박혜경, 이현주, 김은경, 정동일, 팔당호 조류발생 특성 및 수질환경인자의 통계적 분석, *한국물환경학회지*, 21(6), pp. 584-594 (2005).
- 백성욱, 박옥현, 도시대기오염예보모델링을 위한 고유벡터 방법의 적용, *동아대학교 대학원 논문집*, 6, pp. 393 (1982).

- 부산광역시, 낙동강하류·회동수원지 수계 물관리 연구 (2002).
- 신성교, 박청길, 송교육, 주성분 분석법을 이용한 낙동강 하구 해역의 수질 평가, *한국환경과학회지*, 7(2), pp. 171-176 (1998).
- 오영민, 백성욱, 주성분분석을 이용한 하천수질 자료의 해석, *대한환경공학회지*, 6(1), pp. 56-62 (1984).
- 장주영, 다변량 해석에 의한 낙동강 하구 해역 수질특성의 평가, 석사학위논문, 부경대학교 (1999).
- 환경부, 수질오염공정시험법 (2004).
- Ohta, R., Shima, K. and Takayanagi, T., Water quality analysis of the Mikawa Bay by principal component analysis, *Water and Waste*, 28(11), pp. 1118-1121 (1986).