



## Analysis of water quality characteristics of the Miho River basin using multivariate statistical analysis

Yu, Nayoung<sup>a</sup> · Choi, Byeoul<sup>b</sup> · Seo, Dongil<sup>c\*</sup>

<sup>a</sup>Director, Humanplanet Co., Ltd., Daejeon, Korea

<sup>b</sup>Assistant, Environmental Consulting Dept., Humanplanet Co., Ltd., Daejeon, Korea

<sup>c</sup>Professor, Department of Environmental Engineering, Chungnam University, Daejeon, Korea

Paper number: 24-072

Received: 13 September 2024; Revised: 9 October 2024; Accepted: 10 October 2024

### Abstract

In this study, the Seasonal and Trend Composition Using Loess (STL) technique, which enables time series analysis, was applied to indirectly identify the contribution of water quality between nearby water quality measurement points and to interpret the influence between the main stream and tributary streams in the Miho River basin. This technique can identify and remove outliers in the time series data of the water system and analyze water quality change trends and seasonal characteristics. In addition, in order to analyze the correlation and similarity between water quality data at adjacent measurement points, a correlation analysis was conducted, and the points where common factors affect the water quality of the mainstream and inflow tributaries were classified. As a result of factor analysis between points judged to have high correlations in Miho River, it was found that the influence of the mainstream was absolute over tributaries, and it was analyzed that water quality management for the mainstream including the upstream of the Miho River should be given priority. As the degree of water quality influence between points can be analyzed by the method used in this study, it is judged that it can be effectively applied to the development of alternatives for water quality management of rivers in the future.

**Keywords:** Miho River, Water quality, Multivariate statistical analysis, STL decomposition, Correlation analysis, Factor analysis

## 다변량 통계분석을 이용한 미호강 유역 수질 특성 분석

유나영<sup>a</sup> · 최별<sup>b</sup> · 서동일<sup>c\*</sup>

<sup>a</sup>(주)휴먼플래닛 이사, <sup>b</sup>(주)휴먼플래닛 환경컨설팅부 사원, <sup>c</sup>충남대학교 환경공학과 교수

### 요지

본 연구에서는 인근 수질측정지점 간의 수질의 기여도를 간접적으로 파악하고 분류와 지류하천 간의 영향력을 해석하기 위해 시계열 분석이 가능한 STL (Seasonal and Trend decomposition using Loess) 기법을 수질자료에 적용한 결과를 보고하고 있다. STL 기법은 수계 시계열 자료의 이상치를 판별하고 제거할 수 있게 하며 해당 자료의 수질 변화 추세와 계절적 특성을 분석하는 방법이다. 또한, 인접하게 위치한 수질측정지점자료 간의 수질 상관도 및 유사도를 분석하기 위해 지점간 상관분석을 실시하였으며 요인분석을 통해 분류와 유입 지류 측정 자료 간의 수질 자료에서 공통요인이 영향을 미치는 지점을 분류하였다. 미호강의 경우 상관도가 높다고 판단된 지점 간의 요인분석 결과 지류하천 보다 분류의 영향이 절대적인 것으로 파악됨에 따라 미호강 상류를 포함한 분류지점에 대한 수질관리가 우선시되어야 할 것으로 분석되었다. 본 연구에서 사용된 방법으로 우리나라 주요 하천 분류의 상·하류 지점과 유입지점 간의 수질 영향력 정도를 분석할 수 있음에 따라 향후 하천의 수질관리 대안 개발에 효과적으로 적용할 수 있을 것으로 판단된다.

**핵심용어:** 미호강, 수질, 다변량 분석, STL 분해, 상관분석, 요인분석

\*Corresponding Author. Tel: +82-42-821-6679

E-mail: seodi@cnu.ac.kr (Seo, Dongil)

## 1. 서론

현장의 수질 실측자료는 효과적인 물환경 관리를 위한 필수 조건으로 시공간적 측정 결과를 분석함으로써 수질오염의 원인, 배경 등을 파악할 수 있는 중요한 기초자료이다. 환경부는 관련 법령에 따라 하천·호소 등 공공수역에 대한 수질 현황 파악 및 추세파악을 위해서 물환경측정망을 구축·운영하고 있다. 최초의 물환경측정망은 전국 하천 및 호소 등 공공수역에 대한 수질 및 수생태계 현황을 종합적으로 조사하여 수질 및 수생태계의 변화 추세를 파악하는 현황조사에서 시작하였으나, 이후 그간 집행된 주요 정책 및 예산 사업의 효과분석, 장래 물환경정책 수립의 기초 자료 제공 등 다양한 형태로 확대·발전해오고 있다(ME, 2021).

하천의 수질은 복잡하고 다양한 원인에 따라 변화되기 때문에, 단순히 단변량 분석만으로는 이러한 변화의 원인과 패턴을 충분히 이해하기 어렵다. 다변량 통계기법을 활용하면 여러 수질 측정지점의 상호작용과 관계를 동시에 분석할 수 있어 하천의 수질 상태를 보다 종합적으로 평가할 수 있다(Rotiroti *et al.*, 2019; Vieira, *et al.*, 2012). 또한, 다변량 통계기법은 복잡한 데이터 구조를 단순화하여, 주요 요인을 추출하고 이들 요인이 수질에 미치는 영향을 체계적으로 분석할 수 있는 도구를 제공한다(Alharbawee *et al.*, 2024). 따라서 수질 데이터에 다변량 통계기법을 적용한 연구가 국내외로 다양하게 적용되고 있다. 다변량 분석은 강물 수질에 영향을 미치는 주요 요인을 식별할 수 있게 하며(Giri *et al.*, 2019), 복잡한 수질 자료를 평가하여 오염원과 계절적 변화를 밝혀낼 수 있다(Kilic and Yucel, 2018). 다변량 통계분석을 이용하여 낙동강 분류의 시·공간적 수질 분석을 통해 관리 대상 수질 인자와 그 중요성을 밝혀내어 여름/가을철 수질관리가 전 기간에 걸쳐 지배적인 영향을 줄 수 있음을 확인하였으며(Seo *et al.*, 2019), 상관분석, 주성분 및 요인분석을 이용하여 낙동강 주요 지류의 수질 특성을 분석하여 다변량 통계기법이 수질 관리 및 정책 수립에 유용하게 활용될 수 있음을 시사한 바 있다(Park *et al.*, 2019). 금호강 중·상류 유역의 수질에 대한 다변량 통계분석을 수행한 연구에서는 수질항목 간의 상관성을 확인하고 수질오염도가 높은 소유역을 군집으로 분류하였으며(Seo *et al.*, 2021), 다변량 통계분석을 이용하여 국내 대규모 담수호의 수질을 분석하고 클로로필 a를 예측하였다(Lee *et al.*, 2022). 본 연구의 대상 유역과 같은 미호강 유역에 대한 관련 연구에서는 수질측정망 자료를 이용한 다변량 통계분석을 통하여 미호강 수계의 수질에 가장 큰 영향을 미치는 인자를 질소계 열의 농도로 파악한 바 있으며(Yoon *et al.*, 2019), 지류하천별

배출부하량을 조사하여 미호강 수질개선을 위한 대응방안을 모색한 연구가 있다(Song *et al.*, 2021).

측정망은 운영목적에 따라 조사항목, 측정주기, 측정기관 등이 다르다. 유역 내 발생하는 특정 사건에 의해 발생하는 수질 오염원의 유입은 계절성과 추세에서 벗어나 발생하기 때문에 하천의 경우 체류시간이 짧고, 측정망 측정일이 서로 다르므로 특정 사건에 대한 유하거리별 영향을 파악하기에는 어려움이 있다. 또한 수질 측정 결과는 시험자, 시험방법 및 기기 등에 의한 오차도 존재하기 때문에 각 지점의 수질변화는 주로 시간에 따른 변화, 즉 추세를 기반으로 이루어진다. 따라서 본 연구에서는 불규칙한 수질 데이터를 추세 성분과 나머지 잔차 성분으로 구분하여 각 성분을 기반으로 지점간 연속성 및 상관성 분석을 함으로써 유량 혹은 오염부하량이 부재하더라도 분류와 유입지류간 수질 영향력을 해석할 수 있는 방안을 모색하여 효율적인 수질관리를 위한 기초자료를 마련하고자 하였다.

## 2. 연구대상 및 연구방법

### 2.1 연구 대상유역 및 수질 조사지점

연구대상 유역은 금강유역의 대표 지류인 미호강으로 선정하였으며, 미호강 유역의 유역면적은 1,854.95km<sup>2</sup>, 유역둘레의 길이는 280.07km로 금강 전체 유역면적의 약 20%를 차지한다. 미호강은 음성군 삼성면에 위치한 마이산에서 발원하여 경기도, 충청북도, 충청남도, 세종특별자치시를 지나며, 한천, 백곡천, 무심천 등 여러 지류를 포함하고 있다. 미호강 상류 유역은 축산 및 산업시설 등 오염원 밀집 지역이며, 중·하류 유역은 최근 세종시 등의 개발·확장에 의한 환경 변화가 두드러지는 지역으로 수질 개선 대책이 효율적으로 작용할 수 있는 하천 및 지역 선정에 대한 신중한 검토가 필요한 지역이다. 이에 미호강 유역 내 총 32개의 하천 및 산단하천 측정망을 대상으로 분석하였으며, 측정망은 성격에 따라 금강유역환경청, 금강물환경연구소 및 지자체로 조사기관이 나뉘어 있다. 조사지점은 Fig. 1 및 Table 1과 같으며, 수질은 환경부 물환경정보시스템(WEIS, 2024)에서 제공하는 2013년부터 2022년의 일별 수질측정자료를 이용하였다. 수질 항목은 유기물, 영양염류를 대표하는 주요 수질항목인 생물화학적 산소 요구량(BOD), 화학적 산소요구량(COD), 부유물질(SS), 총질소(T-N), 총인(T-P), 총유기탄소(TOC), 6개 항목을 대상으로 하였다.

Table 2는 10년간의 미호강 유역의 하천수질 자료 분석결과

이다. 평균수질은 BOD 3.5 mg/L, T-P 0.108 mg/L로 하천생활 환경기준 보통(III)등급에 해당하며, 특히 BOD, SS 및 T-P 항

목은 평균값에 비해 상당한 분산이 있고 최대값도 높은 편이다.

2.2 연구방법

2.2.1 STL 분해

STL (Seasonal and Trend decomposition using Loess) 분해는 1990년대 초반에 Cleveland 등이 개발한 시계열 분석에서 매우 유용한 기술 중 하나로서, 시계열 데이터를 추세(Trend), 계절성(Seasonality) 및 잔차(Residual) 성분으로 분해한다. 여기서 LOESS는 Locally Estimated Scatterplot Smoothing의 약자로 비선형 관계를 추정하기 위해 구간적으로 가중치를 부여하여 부드러운 추세를 적합시키는 비모수적 회귀기법이다(Hyndman and Athanasopoulos, 2018).

시계열 데이터(y(t))는 STL를 통해 계절 성분 S(t), 추세 성분 T(t), 잔차 성분 R(t)으로 분해할 수 있다(Fig. 2 and Eq. (1)). 계절 성분은 특정 시간 구간(ex. 일, 주, 월, 분기 등)에 따라 정기적으로 반복되는 변동을 포함하며, LOESS Smoothing을 수행하여 계절적 요소를 추정한다(Eq. (1a)). 추세 성분은 장기적인 경향을 나타내며, 시계열 데이터에서 계절성을 제거한 후에 추정한다(Eq. (1b)). 잔차, 즉 불규칙 성분은 계절 성분과 추세 성분을 제거한 후 남은 데이터의 변동성을 말하며, 노이즈 및 데이터의 수집과정에서 발생한 오류, 즉, 설명되지 않은 변동을 포함한다(Eq. (1c)).

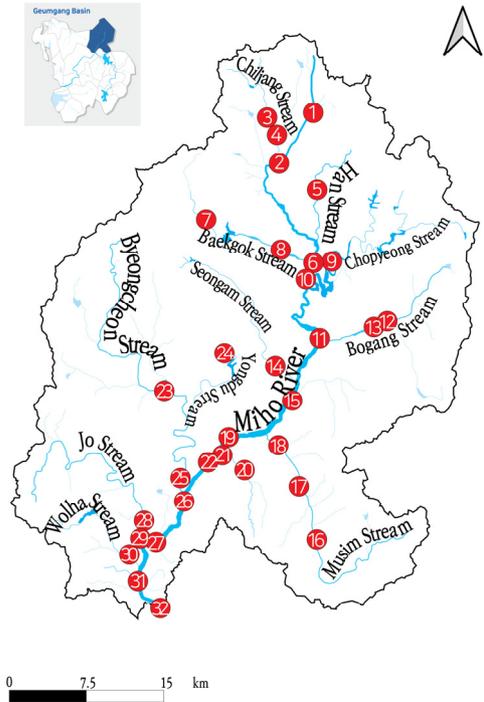


Fig. 1. Location of study area

Table 1. Water quality measurement point of Miho River

No.	Site	No.	Site	No.	Site
1	Mihogang1	12	Bogangcheon1	23	Byeongcheoncheon1
2	Mihogang2	13	Bogangcheon2	24	Yongducheon
3	Chiljangcheon	14	Seong-amcheon	25	Byeongcheoncheon2
4	cChiljangcheon	15	Mihogang6	26	Mihogang8
5	Hancheon	16	Musimcheon1	27	Mihogang9
6	Mihogang3	17	Musimcheon2	28	Jocheon1
7	Baekgogcheon1	18	Musimcheon3	29	Jocheon2
8	Baekgogcheon2	19	cMihogang	30	Wolhacheon
9	Mihogang4	20	cSeognamcheon	31	cBong-amcheon
10	Chopyeongcheon	21	Seognamcheon	32	Mihogang10
11	Mihogang5	22	Mihogang7		

Table 2. Descriptive statistics of water quality data in the Miho River basin

Classification	BOD	COD	SS	T-N	T-P	TOC
Min	0.2	1.1	0.3	0.4	0.004	0.7
Max	40.5	52.5	628.0	21.6	1.910	19.8
Average	3.5	7.3	17.2	4.9	0.108	4.9
Standard deviation	2.6	3.2	36.2	3.4	0.086	0.1

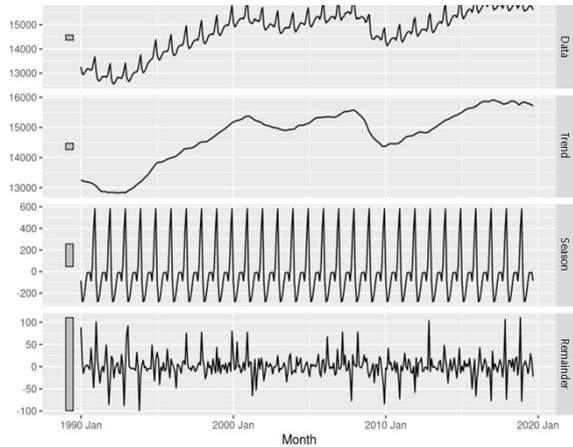


Fig. 2. Components obtained from STL decomposition (trend, seasonal, residual)

$$y(t) = S(t) + T(t) + R(t) \quad (1)$$

$$S(t) = \text{Loess}(y(t)) \quad (1a)$$

$$T(t) = \text{Loess}(y(t) - S(t)) \quad (1b)$$

$$R(t) = y(t) - T(t) - S(t) \quad (1c)$$

STL 분해는 이상값이 존재할 때도 안정적인 분해가 가능하다고 알려져 있다(Cleveland *et al.*, 1990). Oh *et al.* (2017)은 STL 등의 통계기법을 활용하여 서울지역의 PM10 시계열 자료로 미세먼지 특성을 분석하였으며, Lee (2021)은 계절변동이 시간에 따라 변할 때 STL을 활용하여 계절 조정의 유효성을 검토하였다. Rojo *et al.* (2017)은 STL을 사용하여 분해된 데이터 계열의 계절성 및 잔차 요소를 기반으로 공기 중 꽃가루 계열을 예측하기도 했으며, Lafare *et al.* (2016)은 STL을 사용하여 Permo-Triassic Sandstone 대수층의 지하수 거동을 분석한 바 있다. 본 연구에서는 Python 프로그램을 사용하여 시간에 따른 수질데이터를 STL 분해하고 추세 성분과 잔차 성분을 추출하였으며, 각 성분간의 통계분석을 통하여 인근 지점간의 수질 유사도 및 상관도를 분석하고자 하였다.

## 2.2.2 상관분석

상관분석은 두 변수 간의 관계를 통계적으로 분석하는 방법으로, 두 변수 간의 선형적인 관계를 파악하고 한 변수의 변화가 다른 변수와 어떠한 관련이 있는지를 평가한다. 수질 데이터의 시간에 따른 추세에 관한 특성을 분석하기 위해서 두 개의 연속 변수 사이의 선형 관계를 측정하는 피어슨 상관계

수를 선정하여 분석하였다. 본 연구에서는 STL 분해로 얻은 추세 및 잔차 성분에 대하여 Python 프로그램으로 도출한 피어슨 상관계수를 사용하여 상관분석을 수행하였으며, 분석 결과를 통해 변수들 간의 선형관계를 확인하였다. 피어슨 상관계수는 두 변수간의 선형적 관계의 강도와 방향을 나타내며, -1은 완전한 음의 선형관계를, 1은 완전한 양의 선형 관계를, 그리고 0은 선형적 관계가 없음을 의미한다. 피어슨 상관계수는 Eq. (2)로 산정할 수 있다.

Correlation coefficient,  $r$

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (2)$$

$X_i$  = value of the x-variable in a sample

$\bar{X}$  = mean of the value of the x-variable

$Y_i$  = value of the y-variable in a sample

$\bar{Y}$  = mean of the value of the y-variable

## 2.2.3 요인분석

요인분석은 다변량 통계기법 중 하나로 변수들 간의 상호 연관성을 분석하여, 이들 변수들이 공유하는 내재된 공통 요인(Common Factor)을 추출하고, 데이터의 차원을 축소하는 통계 기법이다. 요인분석의 주요 목적은 상관관계가 높은 변수들을 공통 요인 기준으로 묶어, 보다 단순화된 요인 구조를 통해 데이터를 해석하는 데 있다.

본 연구에서는 R 프로그램을 사용하여 요인분석을 수행하였으며, 요인추출은 주성분 분석(Principal Component Analysis, PCA)을 사용하였고 추출된 요인의 회전 방법은 직교(Orthogonal) 회전법 중 Varimax 회전법을 사용하였다. 주성분 분석은 다수의 변수에서 분산을 최대한 설명하는 소수의 요인을 찾아내는 방법으로, 데이터의 차원을 축소하면서도 중요한 정보를 최대한 유지하는 것을 목적으로 한다. 직교 회전법 중에서도 Varimax 회전법은 요인적재량(factor loading)을 최대화하여 결과 해석을 단순화하는 데 효과적이다. 이러한 방법을 통해 다수의 변수간에 공통 요인을 도출하고, 이는 데이터의 복잡성을 줄이고 해석을 용이하게 하는 역할을 한다. 각 변수와 요인 간의 상관관계 정도는 요인적재량으로 나타나며, -1에서 +1 사이의 값으로 표시한다. 요인적재량의 절대적인 기준은 없으나 일반적으로  $\pm 0.4$  이상이면 유의한 변수로 간주하며,  $\pm 0.5$  이상이면 변수와 요인 사이에 적당한 상관관계가 있음을 나타낸다(Hair *et al.*, 2010). 본 연구에서는 강

한 상관관계를 파악하기 위해 요인적재량이  $\pm 0.7$  이상일 때 유의하다고 판단하였다.

### 3. 연구 결과

#### 3.1 STL 분해

환경부 물환경정보시스템의 미호강 지점별 수질(2013~2022년)을 STL 기법을 활용하여 추세 성분과 잔차 성분으로 분해하였으며, 그중 주요 지점의 BOD, T-N, T-P 농도에 대한 STL 결과를 Fig. 3에 표현하였다. 미호강3-미호강4는 모든 수질항목에서 전 기간에 걸쳐 유사한 추세가 나타났으며 잔차 성분 또한 여러 기간에서 일치하는 것으로 나타났다. 미호강6-미호강7의 추세 성분은 2015년의 BOD 및 T-P 외 대부분의 기간에서 유사하게 나타났으나, 잔차 성분은 범위는 유사하지만 이상치가 다르게 나타나는 경향을 보였다. 미호강9-미호강10의 추세와 잔차 성분은 T-N은 유사하게 나타났지만 BOD 및 T-P는 일부 차이를 나타냈다.

#### 3.2 다변량 통계 분석

미호강 유역의 측정지점 간의 수질 연속성 및 유사성을 판단하기 위해서 각 측정항 수질의 STL 분해 결과 중 추세 성분과 잔차 성분을 활용하여 전 지점에 대한 수질항목별(BOD, T-N, T-P, COD, TOC, SS) 상관분석을 수행하였다. 두 지점에서 추세성분의 상관계수가 높게 나타난 경우, 두 지점은 유사한 환경 변화 또는 영향 요인이 작용하고 있다는 것을 나타내어 두 지점의 시간에 따른 수질 변화가 유사한 경향을 나타내는 것을 의미한다. 또한, 잔차 성분의 상관계수가 높게 나타난 경우, 두 지점은 예상치 못한 사건 및 환경 변동 등의 영향이 서로 유사하게 나타내고 있다고 판단하였다.

미호강 유역의 연속 지점에 대한 추세 및 잔차 성분의 상관계수 값을 각각 도출한 결과 주요항목에서 상관계수가 0.7 이상인 지점은 미호강3과 미호강4, 미호강6과 미호강7, 미호강9과 미호강10 지점이었으며, 이 지점들의 상하류 지점간의 수질 연속성이 있다고 판단하고 각 유입지류와 함께 요인분석을 실시하였다.

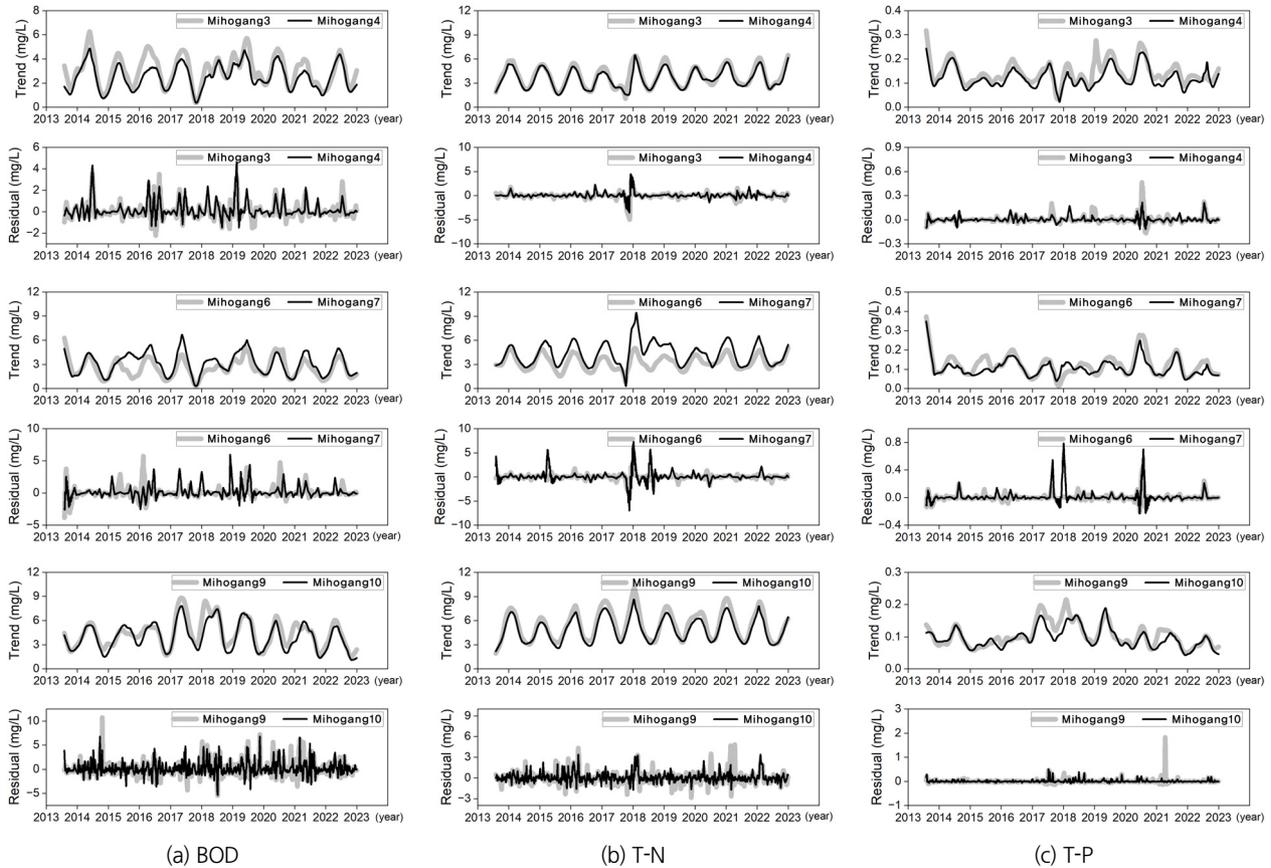


Fig. 3. STL decomposition results (BOD, T-N, T-P)

3.2.1 미호강3·4 및 백곡천2

미호강4 지점의 수질은 상류인 미호강3과 유입지류인 백곡천2 수질의 영향을 받는 지점으로 미호강4에 미치는 상류 수질 영향을 파악하고자 하였다. 백곡천2 주요 수질(BOD, T-N, T-P)은 미호강3·4 수질보다 낮은 수준이고 미호강4의 수질이 미호강3보다 낮다. 수질 농도 범위 및 사분위 범위가 백곡천2보다 미호강3·4에서 비교적 유사한 것으로 나타났다 (Fig. 4 and Table 3).

상·하류 지점간의 수질 유의도 및 연계성을 판단하기 위해 상류 지점인 미호강3, 백곡천2와 미호강4 지점간의 수질 항목

별 상관분석 및 요인분석 결과를 Table 4과 Table 5에 각각 제시하였다. 미호강3-미호강4의 수질은 백곡천2-미호강4 보다 추세 및 잔차 성분의 상관계수가 0.7 이상으로 높게 나타나 지류하천보다 본류의 영향이 더 크다고 해석할 수 있다. 특히 BOD, COD, TOC 및 T-P는 추세와 잔차 성분 모두 상관계수가 높게 나타나 유기물질과 영양염류 유입이 수질에 영향을 미치는 것으로 나타났다(Gwak and Kim, 2016). 요인분석 결과에서는 미호강3-미호강4에서 SS를 제외한 모든 항목의 인적재량이 0.7 이상으로 나타나 해당 지점간의 각 수질은 공통요인에 의해 영향을 받는 것으로 분석된다. 또한 요인1에

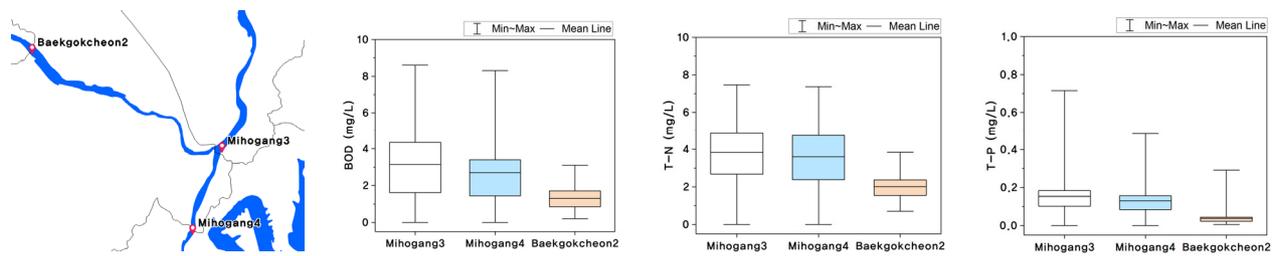


Fig. 4. Locations and Box plots of Mihogangr3-4 and Baekgokcheon2 (13 to '22)

Table 3. Major water quality range of Mihogang3-4 and Baekgokcheon2 (13 to '22)

Site	BOD			T-N			T-P		
	min	max	average	min	max	average	min	max	average
Mihogang3	0.7	8.6	3.1	1.215	7.464	3.927	0.066	0.716	0.158
Mihogang4	0.5	8.3	2.7	1.295	7.375	3.696	0.046	0.485	0.134
Baekgokcheon2	0.2	3.1	1.3	0.696	3.845	1.994	0.004	0.291	0.036

Table 4. Correlation analysis results of Mihogang3-4 and Baekgokcheon2

Site	BOD		T-N		T-P		COD		TOC		SS	
	Trend	Residual	Trend	Residual	Trend	Residual	Trend	Residual	Trend	Residual	Trend	Residual
Baekgokcheon2-Mihogang4	0.49**	0.06**	<b>0.75**</b>	0.04	0.16**	-0.01**	0.54**	0.01**	0.28**	-0.05**	0.14**	0.18**
Mihogang3-Mihogang4	<b>0.93**</b>	<b>0.83**</b>	<b>0.96**</b>	0.56**	<b>0.88**</b>	<b>0.70</b>	<b>0.96**</b>	<b>0.82</b>	<b>0.95**</b>	<b>0.78</b>	<b>0.85**</b>	0.59**

Table 5. Factor analysis results of Mihogang3-4 and Baekgokcheon2

Classification	Factor1			Factor2		
	Mihogang3	Mihogang4	Baekgokcheon2	Mihogang3	Mihogang4	Baekgokcheon2
BOD	<b>0.91</b>	<b>0.86</b>	0.23	0.13	0.11	
T-N			-0.17	<b>0.98</b>	<b>0.99</b>	0.55
T-P	<b>0.71</b>	<b>0.72</b>	0.20	0.15		0.11
COD	<b>0.94</b>	<b>0.94</b>	0.37	-0.10		-0.20
TOC	<b>0.84</b>	<b>0.83</b>	0.38			-0.24
SS	0.49	0.66		-0.10	-0.22	-0.12

의해 유기물 관련 항목(BOD, COD, TOC)과 T-P가 영향을 받는 반면 T-N은 요인2에 의해 영향을 받아 서로 다른 인자에 의해 거동됨을 설명할 수 있다.

3.2.2 미호강6·7, 무심천3, C미호강 및 석남천

미호강6과 하류 미호강7 사이에는 무심천과 석남천이 유입되며 사이에는 산단측정망인 C미호강 지점이 위치해 있다. 본류인 미호강6 및 미호강7의 경우 주요 수질항목에서 평균이 유사한 수준이었으며, 특히 본류 지점의 BOD 항목은 사분위 범위까지 유사하였으나 유입지류(무심천3, 석남천)와의 차이는 크게 나타났다(Fig. 5 and Table 6).

미호강6, 미호강7, C미호강, 무심천3 및 석남천 지점간의 수질 항목별 상관분석 및 요인분석 결과를 Table 7과 Table 8에 각각 제시하였다. 인접 지점간의 상관분석 결과, 지류인 무심천3, 석남천 말단 지점과 본류인 미호강7 지점과의 추세 및 잔차 성분 상관관계수 값이 0.7 미만으로 상관도가 낮은 것으로 나타났으나, 본류인 미호강6-미호강7 및 C미호강-미호강7의 추세 성분 상관관계수 값은 0.7 이상인 것으로 나타나 지류하천보다 본류의 영향이 더 크다고 해석할 수 있다. 특히 COD와 TOC의 상관관계수가 높은 것으로 보아 해당 지역의 산업폐수가 본류에 큰 영향을 미치는 것으로 판단된다. 요인분석 결과에서도 미호강6·7 및 C미호강에서 TOC 항목의 요인적재량이 0.8 이상으로 나타나 산업폐수라는 공통요인이 미치는 영향을 뒷받침할 수 있다. 또한 미호강6·7의 T-P와 SS도

같은 요인에 의해 영향 받는 것으로 보이며, 이는 강우에 의한 오염원 유입이 원인으로 추정된다. 반면 무심천3의 경우 청주시 생활하수 오염원으로 인해 수질오염도가 높고(Kim *et al.*, 2014; Yoon *et al.*, 2019) 분석 기간동안의 BOD와 T-P의 최댓값 또한 높게 나타났지만, 미호강7과의 상관관계수는 낮은 것으로 분석되었다. 이는 추후 실제 유량 및 부하량을 이용하여 지류하천의 기여도와 통합적인 분석을 진행하면 분석의 신뢰도를 높일 수 있을 것으로 사료된다.

3.2.3 미호강9·10, 조천2, 월하천 및 C봉암천

미호강9와 하류 미호강10 사이에는 조천, 월하천 및 봉암천이 유입된다. 본류인 미호강9, 미호강10의 경우 주요 수질항목에서 미호강10의 수질이 약간 낮으나 평균과 범위가 유사한 수준이었으며, C봉암천의 BOD와 미호강9의 T-P의 경우 다른 지점에 비해 편차가 크게 나타났다(Fig. 6 and Table 9).

미호강9, 미호강10, 조천2, 월하천 및 C봉암천 지점간의 수질 항목별 상관분석 및 요인분석 결과를 Table 10와 Table 11에 각각 제시하였다. 인접 지점간의 상관분석 결과, 지류하천보다 본류인 미호강9-미호강10에서 추세 성분의 상관관계수 값은 SS 항목 외 모든 항목에서 0.7 이상인 것으로 나타나 지류하천보다 본류의 영향이 더 크다고 해석할 수 있다. 하지만 미호강9-미호강10의 잔차 성분 상관관계수는 낮은 것으로 보아 일시적인 수질변동은 서로 다른 사건에 의한 것으로 추정되며, 이는 서로 다른 수질 환경이지만 지류하천이 본류에 미치는

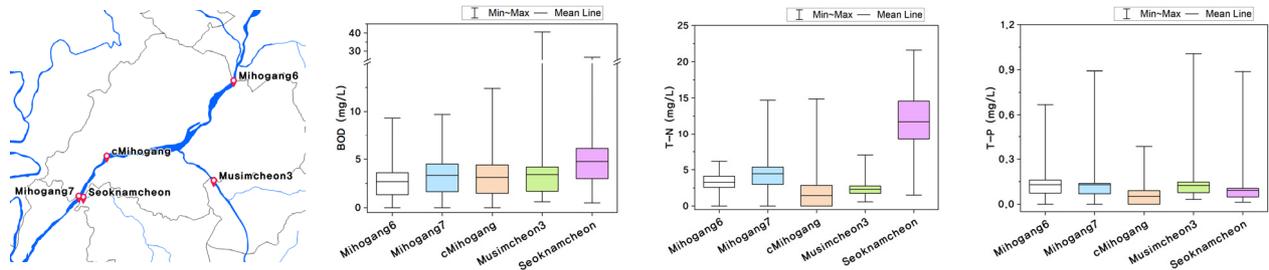


Fig. 5. Locations and Box plots of Mihogang6-7, cMihogang, Musimcheon3 and Seoknamcheon ('13 to '22)

Table 6. Major water quality range of Mihogang6-7, cMihogang, Musimcheon3 and Seoknamcheon ('13 to '22)

Site	BOD			T-N			T-P		
	min	max	average	min	max	average	min	max	average
Mihogang6	0.6	9.3	2.7	1.110	6.161	3.323	0.035	0.668	0.130
Mihogang7	0.5	9.7	3.3	1.405	14.706	4.501	0.034	0.892	0.130
Musimcheon3	0.6	40.5	3.5	0.580	7.033	2.345	0.031	1.006	0.127
cMihogang	0.5	12.4	3.2	0.966	14.896	2.992	0.110	0.384	0.110
Seoknamcheon	0.5	26.5	4.6	1.519	21.594	11.256	0.013	0.890	0.089

**Table 7.** Correlation analysis results of Mihogang6-7, cMihogang, Musimcheon3 and Seoknamcheon

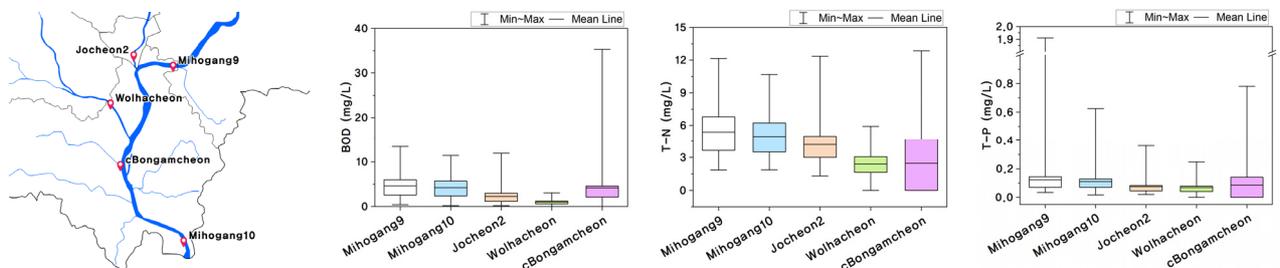
Site	BOD		T-N		T-P		COD		TOC		SS	
	Trend	Residual	Trend	Residual	Trend	Residual	Trend	Residual	Trend	Residual	Trend	Residual
Mihogang6 -cMihogang	<b>0.70**</b>	0.20**	<b>0.86**</b>	0.09**	0.51**	0.06	<b>0.79</b>	0.36	<b>0.91*</b>	0.32	0.37**	0.46
cMihogang -Mihogang7	<b>0.85**</b>	0.26	<b>0.74**</b>	0.08	0.44**	0.03**	<b>0.81**</b>	0.28**	<b>0.84**</b>	0.38	0.37**	0.40
Mihogang6- Mihogang7	<b>0.79**</b>	0.28**	<b>0.80**</b>	0.39**	<b>0.88**</b>	0.56**	<b>0.91**</b>	<b>0.73</b>	<b>0.86**</b>	0.52	<b>0.92**</b>	0.52*
Musimcheon3- Mihogang7	0.03	-0.02**	0.51**	0.04**	0.12**	0.15**	0.38**	0.04**	0.26**	0.01**	0.19**	0.20
Seoknamcheon- Mihogang7	0.31**	0.05**	0.66**	0.05**	0.32**	0.09**	0.57**	0.21**	0.67**	0.12**	0.25**	0.51**

\* p-value <0.05, \*\* p-value <0.01

**Table 8.** Factor analysis results of Mihogang6-7 and cMihogang

Classification	Factor1			Factor2		
	Mihogang6	Mihogang7	cMihogang	Mihogang6	Mihogang7	cMihogang
BOD	0.53	<b>0.76</b>	<b>0.74</b>	0.44		
T-N	-0.21				-0.23	-0.16
T-P	0.21		0.25	<b>0.82</b>	<b>0.64</b>	
COD	0.67	0.64	0.91	0.44	0.41	
TOC	<b>0.84</b>	<b>0.78</b>	<b>0.93</b>			-0.19
SS		-0.11		<b>0.92</b>	<b>0.91</b>	0.59

\* Musimcheon3 and Seoknamcheon were excluded from factor analysis due to their correlation coefficient with Mihogang7 being less than 0.7



**Fig. 6.** Locations and Box plots of Mihogang9-10, Jocheon2, Wolhacheon and cBongamcheon ('13 to '22)

**Table 9.** Major water quality range of Mihogang9-10, Jocheon2, Wolhacheon and cBongamcheon ('13 to '22)

Site	BOD			T-N			T-P		
	min	max	average	min	max	average	min	max	average
Mihogang9	0.4	13.4	4.5	1.854	12.142	5.332	0.034	1.910	0.120
Mihogang10	0.2	11.4	4.2	1.869	10.697	4.907	0.017	0.625	0.109
Jocheon2	0.2	11.9	2.2	1.315	12.359	4.071	0.020	0.362	0.073
Wolhacheon	0.3	3.1	1.0	0.364	5.922	2.486	0.020	0.248	0.072
cBongamcheon	0.8	35.3	4.2	1.510	12.862	4.999	0.027	0.781	0.171

**Table 10.** Correlation analysis results of Mihogang9-10, Jocheon2, Wolhacheon and cBongamcheon

Site	BOD		T-N		T-P		COD		TOC		SS	
	Trend	Residual	Trend	Residual								
Mihogang9 -Mihogang10	<b>0.91**</b>	0.64	<b>0.98**</b>	0.63	<b>0.89**</b>	0.39	<b>0.96**</b>	0.66**	<b>0.96**</b>	<b>0.71</b>	0.53**	0.68
Jocheon2 -Mihogang10	0.67**	0.21**	0.45**	0.14**	0.50**	0.42*	<b>0.85**</b>	0.20	<b>0.89**</b>	0.32	0.51**	0.63**
Wolhacheon -Mihogang10	0.42**	-0.09**	0.46**	0.00**	0.14**	0.09**	0.51**	-0.05**	0.50**	-0.03**	0.17**	0.02**
cBongamcheon -Mihogang10	0.26**	0.09**	0.62**	0.04	0.41**	0.07	<b>0.76**</b>	0.11**	0.58**	0.02	0.28**	0.14**

\* p-value <0.05, \*\* p-value <0.01

**Table 11.** Factor analysis results of Mihogang9-10, Jocheon2, Wolhacheon and cBongamcheon

Classification	Factor1					Factor2				
	Miho-gang9	Miho-gang10	Wolha-cheon	Jocheon2	cBongam-cheon	Miho-gang9	Miho-gang10	Wolha-cheon	Jocheon2	cBongam-cheon
BOD	<b>0.87</b>	<b>0.88</b>	0.22	0.54	0.30	-0.18		0.21		
T-N			-0.38	0.38		<b>-0.99</b>	<b>-0.96</b>	-0.24	-0.41	-0.31
T-P	0.40	0.46		0.22	0.13		0.21	0.51	0.21	0.15
COD	<b>0.94</b>	<b>0.90</b>	0.29	0.55	0.42		0.12	0.58	0.18	0.32
TOC	<b>0.95</b>	<b>0.92</b>	0.25	0.66	0.37	-0.14		0.30	0.19	0.26
SS	0.29	0.29	-0.12	0.22	0.11	0.34	0.34	0.37	0.27	0.18

영향이 적어 본류 추세만 상관도가 높은 것으로 파악된다. 예외적으로 COD가 지류인 조천2 및 C봉암천과 미호강10의 상관계수가 0.7 이상으로 나타나 각 지류의 산업폐수가 미호강10에 영향을 미치는 것으로 분석된다. 요인분석 결과에서는 미호강9-10에서 유기물 관련 항목(BOD, COD, TOC)의 요인적재량이 0.8 이상으로 나타나 생활하수와 같은 공통요인에 의해 영향을 받는 것으로 분석된다. 또한 유기물 관련 항목(BOD, COD, TOC)은 요인1에 의해 영향을 받는 반면 T-N은 요인2에 의해 음의 요인적재량이 높게 나타난 것으로 보아 해당 요인과 음(-)의 상관관계를 갖고 있음을 설명할 수 있다 (Cho *et al.*, 2017).

#### 4. 결론

측정시기가 불규칙한 시계열 데이터인 수질 자료를 STL 분해하여 추세 성분과 잔차 성분을 추출하였으며, 수질 자료의 추세 성분은 경향 및 추세, 잔차 성분은 예상치 못한 오염사건, 측정 오류 등으로 설명할 수 있다. 본 연구는 추출한 성분으로 인접 수질 간의 상관분석 및 요인분석을 실시하였으며, 분석결과로 인접 측정망 간의 수질 연속성을 파악하고 지류 하천의 영향력을 분석하였다. 상관분석에서 추세 성분의 상관

계수가 높은 경우에는 해당 지점의 수질변동 추세가 유사하다고 판단하였으며, 잔차 성분의 상관계수가 높은 경우에는 동일 사건 및 환경이 해당 지점의 수질에 영향을 미친다고 해석하였다. 또한, 요인분석에서 요인적재량이 높게 나타난 경우, 두 지점의 해당 수질항목이 공통된 요인에 의해 영향을 받는다고 분석하였다.

미호강 유역의 측정지점 간 수질 연속성 및 유사성을 판단하기 위해서 32개 지점 간의 주요 수질항목별 상관분석 결과, 본류 지점인 미호강3·4, 미호강6·7, 미호강9·10 지점간의 상관계수가 0.7 이상으로 나타났다. 해당 지점은 주요 지류하천이 유입되는 지점이지만 지류하천의 영향력이 미호강 본류 수질 간의 상관도를 낮출만큼 크지 않다는 것을 의미한다. 또한 요인분석 결과, 미호강3·4의 경우 SS 외 모든 수질 항목, 미호강6·7은 TOC, 미호강9·10 지점은 유기물 관련 수질항목(BOD, COD, TOC)에서 0.7 이상의 높은 상관계수를 나타냈으며, 해당 항목이 공통요인의 영향을 받는 것으로 나타났다. 분석 지점들은 모두 지류하천보다 본류 지점 간의 수질 연속성이 높게 나타났으며, 이는 대부분의 유입지류보다 미호강 본류 농도의 영향이 크다는 연구 결과와 유사하다(Song *et al.*, 2021). 또한 요인적재량이 높았던 수질항목으로 미루어 미호강3·4 수질은 축산 및 농업비점오염, 미호강6·7은 산업폐수,

미호강9·10은 도심하천 및 산업단지 등에서 유입되는 유기물을 해당 구간에 영향을 주는 공통 오염원으로 추정할 수 있다 (Legesse *et al.*, 2022; Rim, 1999). 향후 해당 측정망 영향 범위에서 공통 요인으로 추정되는 오염원을 세부적으로 파악하고 분석한다면 수질 저감을 위한 적절한 관리대책을 마련할 수 있을 것이다.

본 연구에서는 수질자료와 같이 불규칙한 시계열 데이터를 STL 기법을 사용하여 분해함으로써 추세·계절·잔차 성분을 목적에 맞게 적용할 수 있음을 확인하였으며, 향후 수질이 상치를 제거하기 위한 전처리 작업이나 계절에 따른 수질 변화 분석시에도 활용 가능할 것으로 판단된다. 또한 지점간 다변량 통계기법을 적용함으로써 유량 혹은 오염부하량 부재시 인근 지점 간의 수질의 기여도를 간접적으로 파악하고 본류와 지류하천 간의 영향력을 해석할 수 있는 과학적인 분석이 도출됨에 따라 이를 수질 분석의 전략적 관점에서 활용할 수 있을 것으로 기대된다.

## 감사의 글

본 논문은 금강수계 환경기초조사사업의 지원으로 수행되었습니다.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

- Alharbawee, N.A., and Mohammed, A.J. (2024). "Water quality assessment of tigris river using multivariate statistical techniques." *Iraqi Journal of Science*, Vol. 65, pp. 1266-1275.
- Cho, Y.C., Lee, S.W., Ryu, I.G., and Yu, S.J. (2017). "Assessment of spatiotemporal water quality variation using multivariate statistical techniques: A case study of the Imjin River Basin, Korea." *Journal of Korean Society of Environmental Engineers*, Vol. 39, No. 11, pp. 641-649.
- Cleveland, R.B., Cleveland, W.S., McRae, J.E., and Terpenning, I. (1990). "STL: A seasonal-trend decomposition procedure based on loess." *Journal of Official Statics*, Vol. 6, No. 1, pp. 3-73.
- Giri, A., Bharti, V.K., Kalia, S., Kumar, K., Raj, T., and Chaurasia, O.P. (2019). "Utility of multivariate statistical analysis to identify factors contributing river water quality in two different seasons in cold-arid high-altitude region of Leh-Ladakh, India." *Applied Water Science*, Vol. 9, 26.
- Gwak, B.R., and Kim, I.K. (2016). "Characterization of water quality in Changnyeong-Haman weir section using statistical analyses." *Journal of Korean Society of Environmental Engineers*, Vol. 38, No. 2, pp. 71-78.
- Hair, J.F., Black, W.C., Babin, B.J., and Anderson, R.E. (2010). *Multivariate data analysis*. 7th Edition, Pearson Education International, NJ, U.S.
- Hyndman, R.J., and Athanasopoulos, G. (2018). *Forecasting: Principles and practice, 2nd edition*, OTexts, Melbourne, Australia.
- Kilic, E., and Yucel, N. (2018). "Determination of spatial and temporal changes in water quality at Asi River using multivariate statistical techniques." *Turkish Journal of Fisheries and Aquatic Sciences*, Vol. 19, No. 9, pp. 727-737.
- Kim, J.I., Choi, J.W., and An, K.G. (2014). "Spatial and temporal variations of water quality in an urban Miho Stream and some influences of the tributaries on the water quality." *Journal of Environmental Science International*, Vol. 23, No. 3, pp. 433-445.
- Lafare, A.E., Peach, D.W., and Hughes, A.G. (2016). "Use of seasonal trend decomposition to understand groundwater behaviour in the Permo-Triassic Sandstone aquifer, Eden Valley, UK." *Hydrogeology Journal*, Vol. 24, No. 1, pp. 141-158.
- Lee, D.H., Kang, E.T., Joo, J.C., Go, H.W., Ahn, C.M., Bae, Y.H., and Song, K.D. (2022). "Water quality analysis and Chl-a prediction of 15 large-scale freshwater lakes in Korea by multivariate statistical analysis." *Journal of Korean Society of Environmental Engineers*, Vol. 44, No. 12, pp. 589-602.
- Lee, G.H. (2021). "Seasonal adjustment of Korean time series using STL." *Journal of the Korean Official Statistics*, Vol. 26, No. 2, pp. 31-51.
- Legesse, N.S., Kim, J.Y., and Seo, D.I. (2022). "Evaluation of significant pollutant sources affecting water quality of the Geum River using principal component analysis." *Journal of Korea Water Resources Association*, Vol. 55, No. 8, pp. 577-588.
- Ministry of Environment (ME) (2021). *Basic plan for installation and operation of water environment measurement network*.
- Oh, S.M., Shin, H.S., Shin, Y.S., and Jeong, H.C. (2017). "Forecasting the particulate matter in Seoul using a univariate time series approach." *Journal of The Korean Data Analysis Society*, Vol. 19, No. 5, pp. 2457-2468.
- Park, J.B., Kal, B.S., and Kim, S.M. (2019). "Application of multivariate statistical techniques to analyze the pollution characteristics of major tributaries of the Nakdong River," *Journal of Wetlands Research*, Vol. 21, No. 3, pp. 215-223.
- Rim, C.S. (1999). "Multivariate analysis of water quality data at 14 stations in the Geum-River watershed." *Journal of The Korean Environmental Sciences society*, Vol. 8, No. 3, pp. 331-336.
- Rojo, J., Rivero, R., Romero-Morte, J., Fernández-González, F., and Perez-Badia, R. (2017). "Modeling pollen time series using seasonal-trend decomposition procedure based on LOESS

- smoothing.” *International Journal of Biometeorology*, Vol. 61, No. 2, pp. 335-348.
- Rotiroli, M., Zanotti, C., Fumagalli, L., Taviani, S., Stefania, G.A., Patelli, M., and Leoni, B. (2019). “Multivariate statistical analysis supporting the hydrochemical characterization of groundwater and surface water: A case study in northern Italy.” *Rendiconti Online Della Societa Geologica Italiana*, Vol. 47, pp. 90-96.
- Seo, M.J., Cho, C.D., Im, T.H., Kim, S.H., Yoon, H.J., Kim, Y.S., and Kim, G.H. (2019). “Statistical analysis of the spatio-temporal water quality characteristics of the Nakdong River.” *Journal of Environmental Science International*, Vol. 28, No. 11, pp. 303-320.
- Seo, Y.M., Kwon, K.H., Choi, Y.Y., and Lee, B.J. (2021). “Assessment of water quality characteristics in the middle and upper watershed of the Geumho River using multivariate statistical analysis and watershed environmental model.” *Journal of Korean Society on Water Environment*, Vol. 37, No. 6, pp. 520-530.
- Song, Y.H., Lee, Y.H., Lee, J.G., Park, Y.K., and Kim, K.E. (2021). *Policy analysis and response measures for improving water quality in Mihocheon*, Report 2021-47, Daejeon Sejong Reserch Institute.
- Vieira, J.S., Pires, J.C., Martins, F.G., Vilar, V.J., Boaventura, R.A., and Botelho, C.M. (2012). “Surface water quality assessment of Lis River using multivariate statistical methods.” *Water, Air & Soil Pollution*, Vol. 223, pp. 5549-5561.
- Water Environment Information System (WEIS) (2024). accessed 13 July 2024, <<https://water.nier.go.kr/web/>>.
- Yoon, H.Y., Kim, J.H., Chae, M.H., Cho, Y.H., and Cheon, S.U. (2019). “Assessment of water quality in the Miho Stream using multivariate statistics.” *Journal of Environmental Impact Assessment*, Vol. 28, No. 4, pp. 373-386.