

## 기후변화 시나리오의 기온상승에 따른 낙동강 남세균 발생 예측을 위한 데이터 기반 모델 시뮬레이션

장가연<sup>1a</sup> · 조민경<sup>1b</sup> · 김자연<sup>1c,2,†</sup> · 김상준<sup>1d</sup> · 박힘찬<sup>1e</sup> · 박준홍<sup>1f</sup>

<sup>1</sup>연세대학교 건설환경공학과 · <sup>2</sup>오하이오주립대 환경보건과학과

## Data-driven Model Prediction of Harmful Cyanobacterial Blooms in the Nakdong River in Response to Increased Temperatures Under Climate Change Scenarios

Gayeon Jang<sup>1a</sup> · Minkyong Jo<sup>1b</sup> · Jayun Kim<sup>1c,2,†</sup> · Sangjun Kim<sup>1d</sup> · Himchan Park<sup>1e</sup> · Joonhong Park<sup>1f</sup>

<sup>1</sup>Department of Civil and Environmental Engineering, Yonsei University ·

<sup>2</sup>Division of Environmental Health Sciences, The Ohio State University

(Received 24 January 2024, Revised 15 April 2024, Accepted 29 April 2024)

### Abstract

Harmful cyanobacterial blooms (HCBs) are caused by the rapid proliferation of cyanobacteria and are believed to be exacerbated by climate change. However, the extent to which HCBs will be stimulated in the future due to increased temperature remains uncertain. This study aims to predict the future occurrence of cyanobacteria in the Nakdong River, which has the highest incidence of HCBs in South Korea, based on temperature rise scenarios. Representative Concentration Pathways (RCPs) were used as the basis for these scenarios. Data-driven model simulations were conducted, and out of the four machine learning techniques tested (multiple linear regression, support vector regressor, decision tree, and random forest), the random forest model was selected for its relatively high prediction accuracy. The random forest model was used to predict the occurrence of cyanobacteria. The results of boxplot and time-series analyses showed that under the worst-case scenario (RCP8.5 (2100)), where temperature increases significantly, cyanobacterial abundance across all study areas was greatly stimulated. The study also found that the frequencies of HCB occurrences exceeding certain thresholds (100,000 and 1,000,000 cells/mL) increased under both the best-case scenario (RCP2.6 (2050)) and worst-case scenario (RCP8.5 (2100)). These findings suggest that the frequency of HCB occurrences surpassing a certain threshold level can serve as a useful diagnostic indicator of vulnerability to temperature increases caused by climate change. Additionally, this study highlights that water bodies currently susceptible to HCBs are likely to become even more vulnerable with climate change compared to those that are currently less susceptible.

**Key words** : Climate change scenarios, Cyanobacteria, Harmful algal bloom (HAB), Machine learning, Nakdong river

<sup>1a</sup> 학부생(Undergraduate Student), [globel@yonsei.ac.kr](mailto:globel@yonsei.ac.kr), <https://orcid.org/0009-0000-6567-3677>

<sup>1b</sup> 학부생(Undergraduate Student), [zeropoirot@yonsei.ac.kr](mailto:zeropoirot@yonsei.ac.kr), <https://orcid.org/0009-0009-0780-6983>

<sup>1c,2</sup> Corresponding author, 박사수료(Ph.D. Candidate), [naturekim@yonsei.ac.kr](mailto:naturekim@yonsei.ac.kr), <https://orcid.org/0000-0002-6512-5790>

<sup>1d</sup> 학부생(Undergraduate Student), [steven1105@yonsei.ac.kr](mailto:steven1105@yonsei.ac.kr), <https://orcid.org/0009-0002-5989-024X>

<sup>1e</sup> 학부생(Undergraduate Student), [2018144032@yonsei.ac.kr](mailto:2018144032@yonsei.ac.kr), <https://orcid.org/0009-0007-5883-5886>

<sup>1f</sup> 교수(Professor), [parkj@yonsei.ac.kr](mailto:parkj@yonsei.ac.kr), <https://orcid.org/0000-0003-1269-4428>

## 1. Introduction

유해 녹조(harmful algal blooms)는 전 세계 담수 생태계에서 증가하고 있으며 국제적으로 환경 난제로 여겨지고 있다(Huisman et al., 2018). 낙동강 녹조 특히 남세균(cyanobacteria) 발생 역시 국내의 환경 난제이다. 낙동강은 영남지역의 최대 수자원 중 하나로, 약 1,000만 인구의 식수 원임에도 불구하고 유역 내 강우 시 미처리된 하수 및 폐수의 유입이나 농업지역의 영양염류 오염의 유입으로 수질관리가 중요한 지역이다(Woo et al., 2020). 기후변화가 가속화되면서 기온상승에 의한 녹조 발생의 정도와 빈도가 증가할 것으로 추정되고 있다. 특히, 폭염으로 인해 수온이 높아지면 수체의 혼합 작용을 감소시키는 성층화가 유발되어 녹조 발생이 촉진되기도 한다(Jöhnk et al., 2008, Paerl and Huisman, 2008). 우리나라 역시 이상기후의 영향으로 폭염 일수의 증가가 관찰되며(National Disaster Management Research Institute, 2023), 4대강 사업 이후에 낙동강에서 남세균 발생으로 상수원 및 친수공간 경제 발령 일수가 지속적으로 증가하는 추세이다. 2022년에는 낙동강 지역의 남세균 창궐로 상수원의 남세균 독소와 수변공간 독소 에어로졸 이슈에 관해 정부와 시민단체 간에 갈등이 심화되면서 사회정치적인 이슈화가 되기도 하였다.

남세균 발생에 의한 맛 및 냄새 물질과 마이크로시스틴(microcystins)과 같은 인체 독소 물질들은 고도정수처리로 제거가 기술적으로 가능하고 낙동강 본류에서 지표수를 직접 취수하는 정수처리장들은 모두 고도정수처리시설을 보유하고 있다. 하지만 2014년 미국 이리호의 대규모 남세균 증식에 의해서 정수된 수돗물 내 조류독소가 검출되어서 음용수 공급 중단 사태로 주민 이주 및 경제적 손실을 초래한 오하이오주 털리도시의 사례로 인해서 수돗물의 조류독소에 대해서 안전하다고 단언할 수는 없다. 당시 이리호의 남세균 과다증식으로 인해 남세균 마이크로시스틴의 농도가  $10\mu\text{g/L}$ 를 초과하였는데(Cho et al., 2020), 이는 대략 50,000 cells/mL의 남세균 수로 국내에서도 관찰되는 수준이기에 미국의 털리도 사태와 같은 일이 우리나라에도 발생할 가능성을 배제할 수는 없다. 더구나 기후변화에 따라서 상수원의 녹조 발생과 수질 악화가 예측되는데, 증가한 오염부하를 저감하기 위해서 정수처리 과정에서 발암물질인 트리할로메탄(trihalomethane, THM)과 같은 소독부산물 발생이 증가할 수도 있다(Bernat-Quesada et al., 2020). 더불어 친수구역에서 조류발생은 조류독소 위해성으로 인해 수영, 낚시, 캠핑 등의 친수활동이 제약되어 지역 경제와 주민 활동에 피해를 주고 있다. 이러한 기후변화에 의해서 남세균 발생에 취약한 지역을 파악하고 상수원 수질, 정수처리시설, 친수공간 환경위해성 등의 관리 강화를 위한 선제적 대책을 마련하는 것이 필요하다. 이를 위한 첫 번째 단계는 신뢰성 있는 예측 모델을 통해 기후변화를 포함한 다양한 미래 환경변화 시나리오에 대한 남세균 발생 영향을 과학적으로 예측하는 것이다.

해당 지역의 조류 발생을 기작 기반으로 예측하기 위해서는 물질수지와 공정(process) 기반의 모형이 주로 활용된다.

물질수지 및 공정 기반 모형은 해당 유역 내 물질수지 정보가 충분해야 하고 단위 기작 공정별로 모형 매개변수 값들이 현지 조건을 대표하도록 획득이 되어야 하는데, 만약 그렇지 않으면 예측 신뢰도가 낮아지는 한계가 있다(Li et al., 2013; Xie et al., 2012). 이러한 공정 기반 모형의 단점으로 인해서 특정 지역의 남세균 발생을 지역의 수질, 기상, 유량 및 유속 등의 자료들로 통계 기법이나 기계 학습 기법을 학습시켜서 예측하는 데이터 기반 모형이 대안으로 최근에 고려되는 추세이다(Kim, Cho et al., 2023; Villanueva et al., 2023). 환경부 물환경정보시스템의 데이터를 이용해서 낙동강 지역의 남세균 수 예측을 위한 데이터 기반의 모형이 국립환경과학원의 공정 기반 모형보다 예측 정확도를 향상한 연구 사례도 보고된 바 있다(Kim, Jung et al., 2023). 남세균 변성의 원인에 여러 요인이 있지만, 기온 및 수온이 대체로 남세균 변성에 가장 큰 영향을 미친다(Cha et al., 2017; Kim et al., 2024). 이러한 연구 동향에 근거하여 데이터 기반의 기계학습(machine learning) 기법을 활용하여 기후변화 시나리오에 따른 미래의 기온 상승이 낙동강 지역 남세균 발생에 미칠 영향을 예측해 보는 연구가 필요하다.

다양한 자연재해 즉 홍수, 가뭄, 해수면 상승, 산사태의 사례들이 기후변화에 따라 그 빈도나 정도가 점점 증가하고 있고(Ministry of Environment, 2020), 자연재해에 취약하거나 피해 이력이 잦은 지역 혹은 시설이 기후변화에 취약할 수도 있다(Bae et al., 2015). 자연재해의 속성을 지닌 남세균과 같은 녹조의 경우도 그 발생 빈도가 높고 녹조 세포의 수가 많을수록 기후변화에 취약할 것으로 추측된다. 하지만 이를 데이터 기반으로 보인 사례가 없으므로, 기후변화 시나리오별로 남세균 수와 발생 빈도를 예측하여 현재 남세균 발생에 대한 지역별 취약성에 따라 기후변화에 더 취약함을 보이는 체계적인 시뮬레이션 연구가 필요하다. 상술한 연구 배경과 필요성에 따라서 국내 녹조 발생의 주요 지역인 낙동강에서 기후변화 시나리오의 기온상승에 따른 미래 남세균 발생을 예측하는 것을 본 연구의 목표로 설정하였다. 본 연구에서는 낙동강 유역 상수원 구간 내에서도 서로 다른 4개 지점을 연구 대상으로 선정하여 기후변화 시나리오인 대표농도경로(Representative Concentration Pathways, RCPs) 중 최상과 최악의 온실가스 배출 시나리오에 해당하는 기온상승에 따른 남세균 발생 예측을 기계학습 기반한 모형 시뮬레이션을 통해서 수행하였다. 본 연구는 기후변화에 따라 미래에 증가할 남세균의 정량을 추정하고 취약한 지역을 미리 파악하여 취약성에 미리 대비할 수 있는 상수원 관리 및 수처리장 운영 전략 수립에 활용할 수 있을 것이다.

## 2. Materials and Methods

### 2.1 연구 대상 지점과 데이터 수집

한국은 수돗물을 안정적으로 공급하고 친수활동 시 조류독소로부터 국민의 안전을 도모하기 위해 1998년부터 조류정보제를 운영 중이다. 물환경보전법에 따라 남세균 수치(cells/mL)를 주기적으로 모니터링하여 일정 수준(관심:

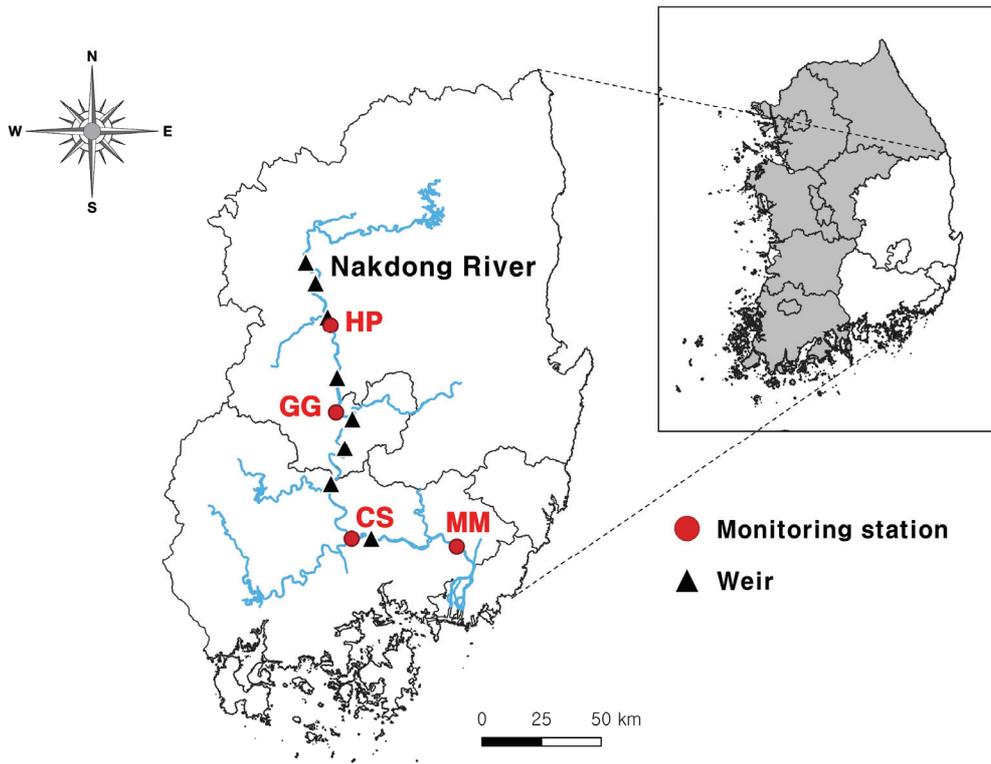


Fig. 1. Map of the study area, showing the four sites (HP, GG, CS, and MM) along the Nakdong River in South Korea.

1,000(상수원구간), 20,000(친수활동구간); 경계: 10,000(상수원구간), 100,000(친수활동구간); 대발생: 1,000,000(상수원구간) 이상의 조류가 발생하면 경보를 발령한다. 본 연구는 국내에서 조류경보 발령일수가 가장 많은(2022년 낙동강 조류경보 발령일수 700일, 전국 조류경보 발령 778일) 낙동강 본류에 위치한 상수원 구간 조류경보지점 4개를 대상으로 연구를 수행하였다(Fig. 1). 상류부터 하류 순서에 따라 해평(HP), 강정·고령(GG), 칠서(CS), 물금·매리(MM)는 모두 상수원으로 사용되며 물금·매리 지점을 제외하고 4대강 사업으로 설치된 보의 위치와 10 km 이내에 있다.

남세균 및 수질 데이터는 물환경정보시스템(<http://water.nier.go.kr>)으로부터 획득하였다. 지점별로 녹조(조류)정보로부터 1주 간격(경계수준 이상이면 1주 2회 이상 측정)의 남세균 측정 데이터를, 수질측정망으로부터 1주 간격의 수질 데이터를 수집했다. 남세균은 하천의 단면 중 수심이 가장 깊은 수면에서 상층(수표면을 포함하여 수심 50 cm 이내)과 수심의 1/3 및 2/3에서 각각 채수한 후(경계수준 이상일 때는 수심이 가장 깊은 수면의 지점과 그 지점을 중심으로 좌우로 수면폭을 2등분한 각각의 지점(좌·중·우)에서 수심별(상·중·하층)로 채수한 후) 혼합하여 현미경 계수법으로 측정되었다. HP, GG, CS 지점에서는 2016~2023년, MM 지점은 2020~2023년 데이터를 수집했다. 수질데이터로는 연구 대상 지점(HP, GG, CS, MM)과 가장 가까운 수질측정망(각각 강정, 성주, 남지, 물금)에서 생물화학적산소요구량(biochemical oxygen demand, BOD), 클로로필-*a*(chlorophyll *a*, Chl-*a*), 화학적 산소요구량(chemical oxygen demand, COD), 용존산소(dissolved

oxygen, DO), 전기전도도(electrical conductivity, EC), 수소이온 농도(pH), 부유물질(suspended solids, SS), 총질소(total nitrogen, TN), 총인(total phosphorus, TP), 수온을 수집했다. 또한, 유속 데이터는 국가수자원관리종합정보시스템(<http://www.wamis.go.kr>)에서 지점별 유량과 유속 관계식을 도출하여 각 지점에서 유량 값을 유속으로 변환하여 수집하였다. 각 지점의 기온은 기상청 기상자료개방포털(<https://data.kma.go.kr>)에 종관기상관측(HP: 구미, GG: 대구, CS: 의령군, MM: 양산시) 자료로부터 수집했다.

## 2.2 기계 학습(machine learning) 기법

기계학습 모형으로는 다중선형회귀(multiple linear regression, MLR), 의사결정나무(decision tree, DT), 서포트벡터회귀(support vector regression, SVR), 랜덤포레스트(random forest, RF)를 사용했다. 이 기법들은 기계학습 알고리즘 중 자주 사용되는 모델이며(Choi et al., 2012; Jung et al., 2021; Jung and Kim, 2016; Moon et al., 2014), 모델 parameter가 무수히 많은 인공신경망을 구축할 때보다 크기가 더 작은 데이터 세트를 활용하기에 적합하다(Aggarwal, 2018). 이 네 가지 모델을 지점별로 학습시켜 가장 좋은 검증 정확도를 보이는 모델로 기후변화에 따른 시뮬레이션을 수행했다. DT와 SVR의 하이퍼파라미터(hyperparameter)는 기본값(DT: criterion= "squared\_error", max\_depth= "None", min\_sample\_split=1 등; SVR: kernel= "rbf", gamma= "scale", C=1 등)으로 설정했고, RF의 하이퍼파라미터 값들은 나무 개수 100개를 두 가지 다른 경우(max\_features= "sqrt", 5)로 나누어 사용했다. 기계 학습 모델링은 Scikit-learn 라이브러리를 활용했고 파이썬(Python) 언어

로 모델 학습과 검증, 예측 및 시뮬레이션을 구현했다.

데이터 전처리, 입력변수 구성과 모델링 방식은 Kim, Jung et al. (2023)과 유사하게 수행했다. 남세균 측정값을 로그변환(log cells/mL)하여 종속변수로, 유사한 일자에 수집된 수질 인자 및 유속을 독립변수(총 11개: BOD, Chl-a, COD, DO, EC, pH, SS, TN, TP, 수온, 유속)로 설정하여 학습데이터로 만들었다. 데이터는 연중 녹조가 주로 발생하는 4~11월에 측정된 남세균 측정값을 사용했으며, 남세균 측정값이 0인 데이터와 독립변수 중 한 개 이상 결측인 데이터를 제거했다. SVR은 입력변수 값의 분산 규모에 영향을 받기에 이 모델을 학습할 때만 모든 독립변수 데이터를 평균이 0, 분산이 1이 되도록 표준화 처리했다.

### 2.3 모델 검증 및 예측 성능 비교 방법

모델 검증은 전체 데이터 중 한 개를 제외한 데이터로 학습하고 나머지 한 개 데이터를 예측하는 과정을 전체 데이터의 개수만큼 반복하는 leave-one-out 교차검증(cross-validation)법을 활용했다. 이를 통해 실측값과 예측값의 상관관계수(correlation coefficient ( $r$ ), 식 1)를 계산하여 예측의 정확도 지표(Xiao et al., 2017)로 보였다. 또한, 추가적인 지표로서 실측값 대비 예측값의 오차를 나타내는 평균제곱오차(mean squared error, MSE, 식 2)를 계산했다. 식 1, 2에서  $y_i$ 는 실측값,  $\bar{y}$ 는  $y$ 의 평균값,  $\hat{y}_i$ 는 예측값을 나타낸다.

$$r = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (\text{식 1})$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (\text{식 2})$$

### 2.4 기후변화 시나리오(RCPs)의 기온상승에 따른 시뮬레이션

대표농도경로(RCPs)란 IPCC 5차 평가보고서에서 온실가스 감축 정도에 따른 기후변화 시나리오이다(IPCC, 2014).

본 연구에서는 최상의 기후변화 시나리오로 RCP2.6, 최악의 시나리오로 RCP8.5를 선정하고, 2가지 기후변화 시나리오에 대해 2가지 시점(2050년, 2100년)에 대해서 각각 시뮬레이션을 수행하였다(Table 1).

기후변화 시나리오에 따른 기온상승 조건을 다음과 같이 활용하였다. 기후변화 시나리오(RCP2.6, RCP8.5)에 따른 미래 월별 기온 데이터를 기상청 기후정보포털(<http://www.climate.go.kr/home/>)에서 지점별로 (HP: 경상북도 구미시, GG: 경상북도 고령군, CS: 경상남도 창원군, MM: 경상남도 김해시) 수집했다. 예측된 미래 기온 데이터의 수집기간은 2041~2050년 및 2091~2100년으로 설정했다. 이를 현재 남세균이 관찰된 기간의 기온과 비교해서 기온 상승분을 계산했다. 기온 상승분을 수온에 반영하기 위해, 지점별 기온과 수온 간 관계식(수온(°C)=0.8806×기온(°C)+4.3015 ( $R^2=0.901$ ))을 도출하여 각 지점에서 상승한 기온을 수온이 상승한 것으로 변환했다. 학습 데이터로 구축한 기계학습 모형에 기후변화 시나리오에 따라 월별 상승한 수온 데이터를 입력하여 예측값을 산출했다. 그 뒤에, 수온 상승에 따른 예측값과 수온 상승 전 기존 모델 출력값의 비례관계를 도출했다. 이 비례관계를 도출할 때, 각 지점의 75퍼센타일 이하의 수온일 때의 남세균 값만 사용했다. 이로써 기계학습 모델이 상대적 소수의 데이터(75퍼센타일을 초과하는 수온 데이터)에 대한 예측 능력이 떨어진다는 한계를 보완하고, 증가한 수온 값이 입력될 때 출력되는 예측값의 신뢰도를 높이고자 했다. 도출한 현재 대비 미래의 비례관계를 사용하여 지점별 현재 남세균 측정값을 기온상승 시나리오에 따른 미래 예측값으로 변환했다.

## 3. Results and Discussion

### 3.1 기계학습 기법별 예측 성능 비교

본 연구에서 사용된 5가지의 기계학습 모델이 남세균 수를 예측하는 성능을 비교하였다(Table 2). 관찰된 실측치와 모형 예측치 간의 상관성을 보여주는 지표인 correlation coefficient의 결과를 보면 RF 기법을 사용한 경우는 대체로 0.7 보다 높은 값들을 보이고 MLR, DT 기법을 사용한 경우에는 보통 0.7 보다 낮은 상관계수를 보였다. 모든 모델에서 실측치와 모형 예측치 간의 상관관계는 유의성을 보였다( $p$ -value < 0.001).

**Table 1.** Definitions of Representative Concentration Pathway (RCP) scenarios used in this study and projected temperatures used for simulations

Representative Concentration Pathway	Definition	Mean air temperature (°C) at the study sites.	
		2041-2050	2091-2100
RCP2.6	CO <sub>2</sub> emissions start declining by 2020 and go to zero by 2100. CO <sub>2</sub> concentration in 2100: 421 ppm. Global temperature increase in 2081-2100 relative to 1985-2005: 0.3-1.7°C.	HP: 13.5 GG: 14.2 CS: 14.5 MM: 15.3	HP: 13.7 GG: 14.4 CS: 14.7 MM: 15.5
RCP8.5	CO <sub>2</sub> emissions continue to rise throughout the 21st century. CO <sub>2</sub> concentration in 2100: 936 ppm. Global temperature increase in 2081-2100 relative to 1985-2005: 2.6-4.8°C.	HP: 13.8 GG: 14.5 CS: 14.7 MM: 15.5	HP: 16.5 GG: 17.1 CS: 17.3 MM: 18.1

**Table 2.** Validation performance comparison of cyanobacterial abundance predictions by different machine learning models.

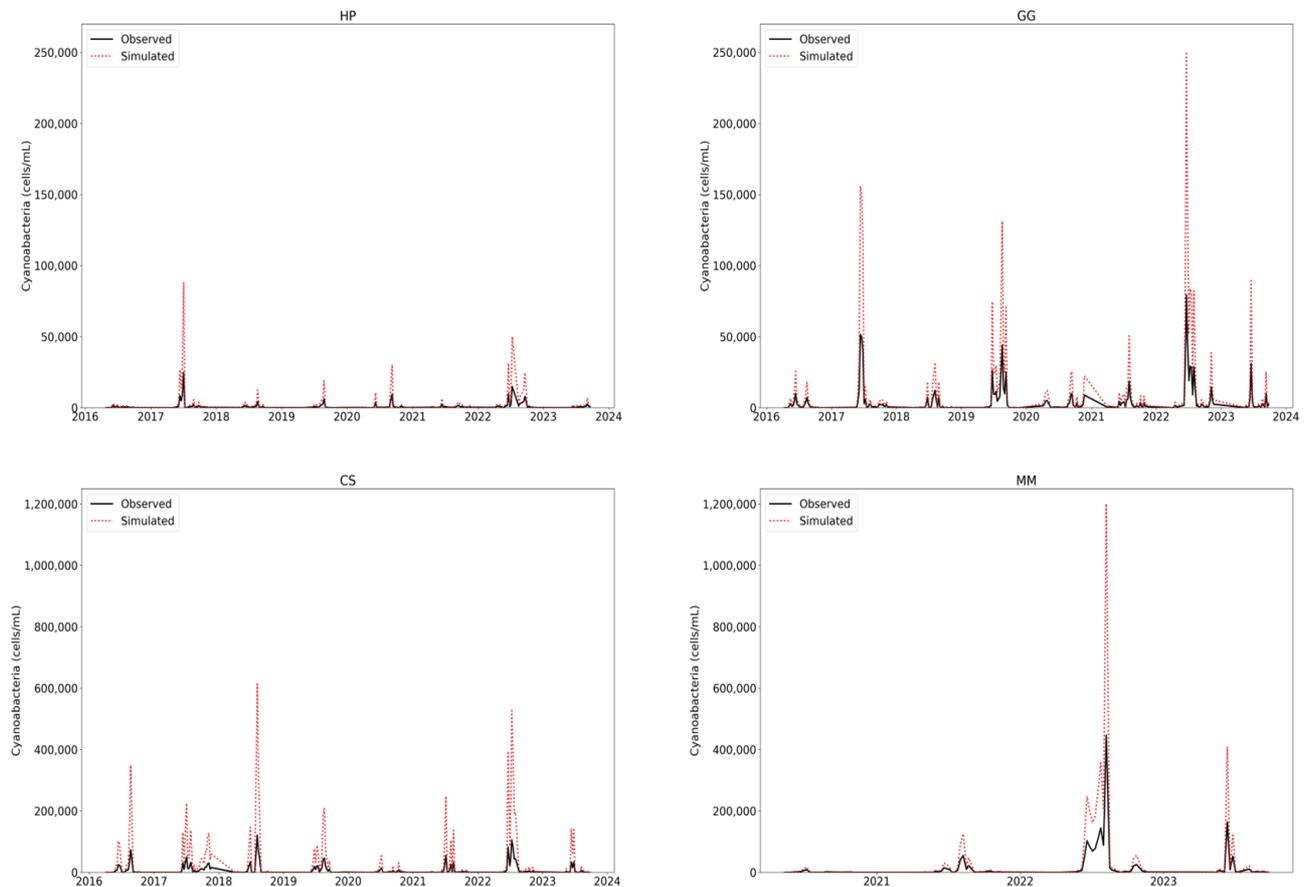
Site	Performance	Multiple linear regression	Support vector regressor	Decision tree	Random forest-“sqrt”	Random forest-5
HP	$r$	0.5081	0.5822	0.3211	0.5989	0.6109
	MSE	0.3057	0.2700	0.5327	0.2610	0.2551
GG	$r$	0.6273	0.6445	0.5584	0.7294	0.7251
	MSE	0.4737	0.4594	0.6538	0.3781	0.3751
CS	$r$	0.7108	0.7406	0.5460	0.7636	0.7583
	MSE	0.4617	0.4200	0.7853	0.3966	0.3997
MM	$r$	0.7615	0.8219	0.6620	0.8181	0.8111
	MSE	0.3806	0.2978	0.5477	0.3121	0.3142

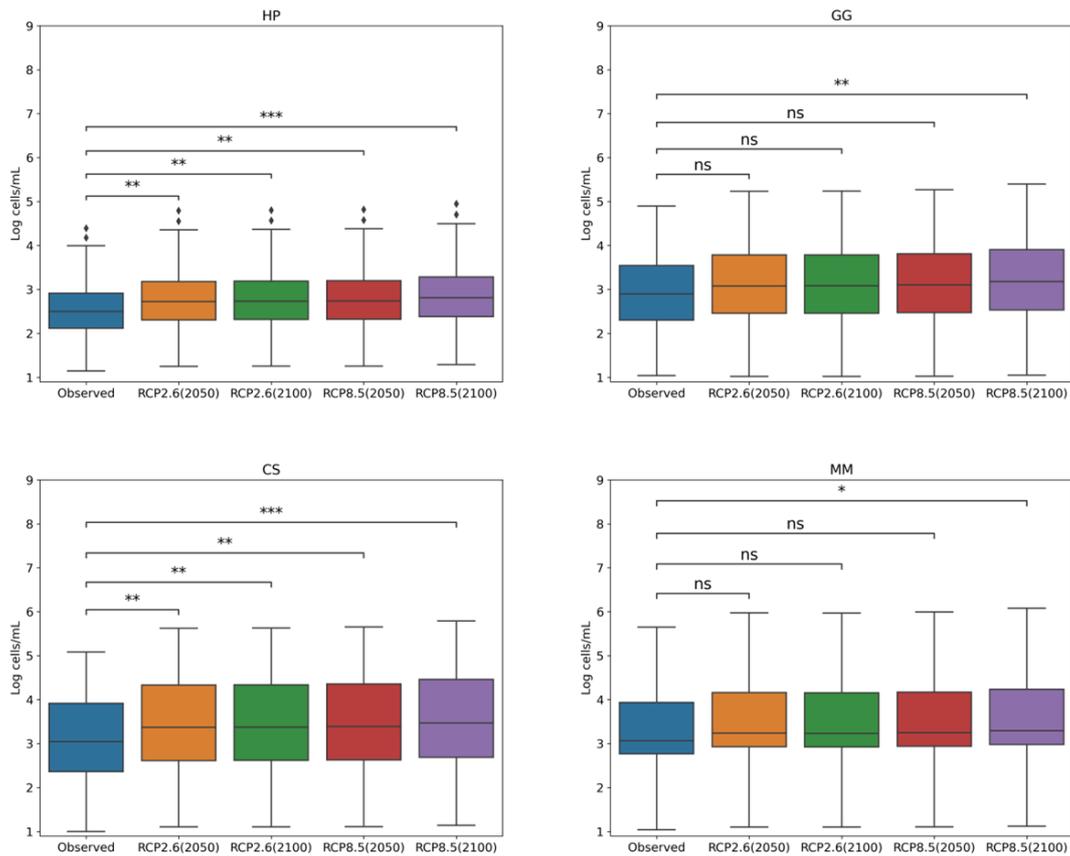
실측치와 예측치 간의 오차를 나타내는 통계지표인 MSE의 값을 보면 일반적으로 RF의 예측 결과가 실측 결과에 통계적으로 근사함을 보였고, SVR의 결과가 두 번째로 근사했다. MLR 및 DT의 예측 정확도는 RF, SVR의 정확도보다 현저히 떨어진 것을 보였다. 따라서 RF 기법이 낙동강 해당 지점들의 남세균 수 예측에 MLR, SVR, DT 기법보다 우수함을 보였고 RF의 절대적인 성능도 준수했다(상관계수 > 0.6). 이에 따라 RCP 기후변화 시나리오에 의한 기온상승에 따른 낙동강 남세균 수 변화 예측을 위한 기계학습 기법으로 RF 기법을 선정하였다. 본 연구에서 사용된 두 가지 다른 RF 기

법들(max features 하이퍼파라미터 설정: “sqrt” 또는 5)은 유사한 예측 성능을 보였다.

### 3.2 기후변화 시나리오 기온상승에 따른 남세균 수 예측 시뮬레이션

최악의 기후변화 시나리오(RCP8.5의 2100년)에서 예측한 남세균 수의 데이터와 현재 관찰된 남세균 수의 데이터를 시계열로 비교해 보았다(Fig. 2). 남세균이 일정 기간(2주 이상) 지속하여 또는 일정 수준 (1,000 cells/mL 이상) 검출된 경우에는 기후변화 시나리오에 의한 기온상승 때문에 남세균 수

**Fig. 2.** Comparison of future simulated cyanobacteria abundances under the worst temperature rise scenario (RCP 8.5 (2100), “simulated”) with current observations (“observed”).



**Fig. 3.** Boxplots of current (“Observed”) and simulated future cyanobacterial abundances under two RCP temperature rise scenarios (RCP2.6 and RCP8.5) for two different prediction times (2050 and 2100). Annotated asterisks indicate statistically significant differences according to p-values obtained from Mann-Whitney U test ( $0.0001 < *** \leq 0.001$ ,  $0.001 < ** \leq 0.01$ ,  $0.01 < * \leq 0.05$ , and “ns” indicates non-significant difference).

가 현격히 증가하는 경향을 보였다. 최악의 기후변화 시나리오에 따라 남세균 수의 최댓값은 HP, GG, CS, MM에서 각각 88,295, 250,506, 613,304, 1,200,578 cells/mL 으로 예측되었다. 낙동강에서 이때까지 관측되지 않았던 대발생 경보 수준 즉, 1 mL 당 100만 세포 수 이상의 수치가 낙동강 하류인 MM에서 발생할 것으로 예측되었다. MM에서는 현재 관측값에 비해 미래 남세균이 최대 약 3배 증가하는 것에 비해 나머지 지점에서는 더 큰 폭인 최대 약 5~6배 증가할 것으로 예측되었다. 따라서, 현재 남세균 검출 수준이 낮은 지점에서가 남세균 검출 수준이 높은 지점보다 기후변화에 의한 기온상승에 남세균 증식이 더욱더 민감하게 반응할 것임을 제시하고 있다.

기후변화 시나리오의 기온상승에 의한 남세균 발생 변화를 더욱 면밀하게 분석하기 위해서 기후변화 시나리오와 예측 시점 조합별로 남세균 수 예측 시뮬레이션 결과값의 분포(4~11월)를 박스 플롯(boxplot)으로 비교해 보았다(Fig. 3). 평균치 결과를 기준으로 보면, 연구 대상의 낙동강 모든 지점에서 현재의 남세균 수 데이터에 비해서 기후변화 시나리오에 따른 기온상승에 대해서 낙동강의 남세균 수는 평균적으로 증가하였고, 이 결과는 예측시점에 무관하게 동일하였다. 반면, 현재의 데이터와 예측된 데이터 간의 통계적 유

의성 분석을 위한 Mann-Whitney test 결과는 연구 대상 지점별로 차이가 있었다. HP 지점과 CS 지점의 경우, 현재의 데이터에 대비 RCP2.6 시나리오의 2050년 예측 결과, RCP2.6 시나리오의 2100년 예측 결과, RCP8.5 시나리오의 2050년 예측 결과에서는  $0.001 < p\text{-value} \leq 0.01$  유의 수준에서 남세균 수가 증가하는 것으로 나타났고, RCP8.5 시나리오의 2100년 예측 결과에서는  $0.0001 < p\text{-value} \leq 0.001$  유의 수준에서 남세균 수가 증가하는 것으로 예측되었다. 즉 HP와 CS 지점에서 기후변화 시나리오의 기온상승에 따라 예측한 미래 남세균 수 증가의 통계적 유의성이 높게 나타났다. GG 지점에서는 현재의 데이터에 대비 RCP2.6 시나리오의 2050년 예측 결과, RCP2.6 시나리오의 2100년 예측 결과, RCP8.5 시나리오의 2050년 예측 결과에서는  $p\text{-value} > 0.05$  유의 수준으로 통계적 유의성이 낮게 나타났고, RCP8.5 시나리오의 2100년 예측 결과는  $0.001 < p\text{-value} \leq 0.01$  유의 수준에서 남세균 수가 증가하는 것으로 예측되었다. GG 지점의 남세균 발생 증가는 HP 지점보다는 통계적인 유의성이 상대적으로 낮지만, 최악의 기후변화 시나리오와 2100년 예측 시점에 해당하는 기온상승에 대해서 남세균 수의 유의한 증가가 예측되었다. 보의 영향을 받는 지점으로 알려진 HP, GG, CS 지점들과 달리 보의 영향이 작은 MM 지점에서는 최악의 시

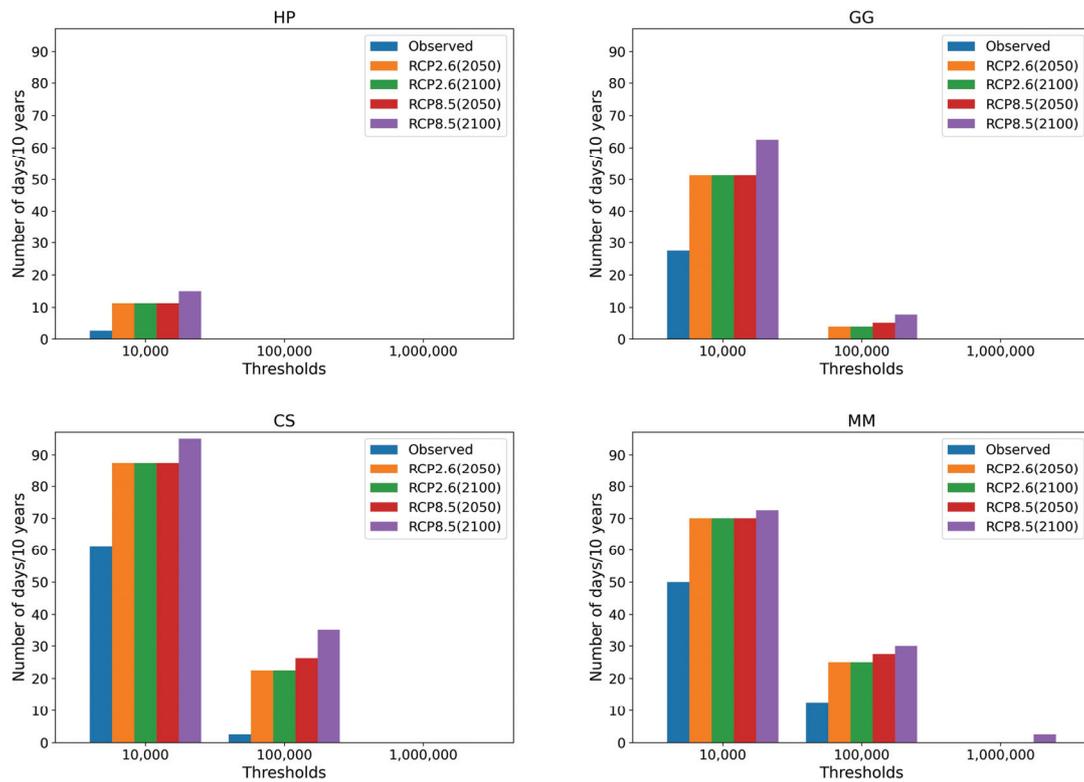


Fig. 4. Barplots showing the number of incidents exceeding the three thresholds of cyanobacterial abundance (cells/mL) currently ("Observed") and under climate change scenarios at the four study sites. The y-axis indicates the number of days over a 10-year period or 521 measurements (i.e., data on a weekly basis).

나리오(RCP8.5의 2100년)를 제외하고 기후변화 시나리오의 기온상승에 따라 예측된 미래 평균 남세균 수의 증가는 통계적으로 유의하지 않았다(p-value > 0.05).

### 3.3 기후변화 시나리오 기온상승에 따른 남세균 발생 빈도에 미치는 영향

남세균 수가 높아질수록 남세균 발생에 의한 상수원 수질과 수변공간 에어로졸과 같은 국민 보건에 미치는 영향이 커질 수 있다(Lee et al., 2022; Wiśniewska et al., 2019). 국내의 경우, 상수원구간에 대한 녹조 현상의 경계, 대발생 수준을 각각 남세균 수가 10,000, 1,000,000 cells/mL로, 친수활동구간의 관심, 경계 수준은 각각 20,000, 100,000 cells/mL로 정하고 있다(Ministry of Environment, 2022). 친수공간에 대해서 캐나다와 세계보건기구(World Health Organization, WHO)에서는 100,000 cell/L로 경보 수준을 정하고 있다(Health Canada, 2022; WHO, 2021). 이러한 상수원 및 친수공간 경계와 대발생에 대한 국내의 기준을 초과하는 남세균 발생의 빈도를 기후변화 시나리오별로 예측하였다(Fig. 4). 최선과 최악의 기후변화 시나리오 모두 현재보다는 상수원 및 친수공간 경계 수준을 초과하는 남세균 발생 빈도는 증가함을 보였다. 최악의 기후변화 시나리오인 RCP8.5의 2100년 예측 결과 연구 지점들의 상수원 경계 기준(10,000 cell/mL) 초과 빈도를 비교해 보면, CS(연 9.5주, 현재 대비 연 3.5주 증가) > MM(연 7.2주, 현재 대비 연 2.2주 증가) > GG(연

6.2주, 현재 대비 연 3.5주 증가) > HP(연 1.5주, 현재 대비 연 1.2주 증가) 순이었다. RCP8.5의 2100년 예측 결과 기준으로 지점들의 친수공간 경계 기준(100,000 cell/mL) 초과 빈도를 살펴보면, CS(연 3.5주, 현재 대비 연 3.4주 증가) > MM(연 3.0주, 현재 대비 1.8주 증가) > GG(연 0.8주, 현재 대비 연 0.8주 증가) > HP(연 0.0주, 현재 대비 0.0주) 순이었다. 이러한 결과들은 기후변화에 의해서 상수원 및 친수공간 경계 발령 일수가 증가해서 이에 따른 비용도 증가될 것과 그러한 증가 폭이 지점별로 차이가 있음을 예측하고 있다. CS 지점과 MM 지점의 경계 기준 초과하는 빈도가 기후변화로 미래에 찾아질 가능성이 GG 지점이나 HP 지점에 비해서 큰 것으로 시뮬레이션 결과가 나타났다. 대발생 기준(1,000,000 cell/mL)을 초과하는 경우는 오직 MM 지점에서 최악의 시나리오인 RCP8.5 2100년 예측의 결과에서만 나타났다. HP 지점의 경우 최악의 기후변화에서도 친수공간 경계 기준이나 대발생 기준을 초과하지 않아서 4개 연구 지점 중 기후변화에 대해서 가장 안정적인 결과를 보였다.

기후변화에 의한 미래의 남세균 발생 정도와 빈도의 증가가 보의 영향에 의해서 더욱더 커질 것이라는 가설 설정이 가능하다. 이 가설을 검증하기 위해서 보와의 거리가 상대적으로 가까워 보의 영향이 있다고 여겨지는 지점들(HP, GG, CS)과 보의 영향이 없다고 여겨지는 지점(MM) 간에 결과를 비교해 보았다. 보의 건설로 인한 효과로는 대표적으로 수위 상승, 수리학적 체류시간 증가, 하수처리시설 규제 강화로

인한 총인 유입량 감소가 있다. 비교 결과 보의 영향이 없다고 여겨지는 MM 지점에서 보의 영향이 있는 HP 혹은 GG 지점보다 기후변화의 기온상승에 의해서 경계 기준 초과 빈도가 더 커지고 심지어 대발생도 발생할 수 있음이 예측되었다. 반면 남세균 발생의 정도(남세균 수)는 보의 영향보다는 수온이 높아질수록 커지는 것으로 보인다. 예로서, 최악의 기후변화 시나리오에 따라 예측된 남세균 수의 최댓값은  $HP < GG < CS < MM$  (각각 약 100,000, 250,000, 600,000, 1,200,000 cells/mL) 순이었으며(Fig. 2), 이는 연평균 기온이 낮음에서 높음의 순서와 일치하였다(Table 1). 비록 MM 지점이 보와의 거리가 멀더라도, 다른 상류 지점이 보 건설에 따라 받은 영향과 MM에 대한 보 건설에 따른 영향이 큰 차이가 없을 가능성도 있다. 본 연구에서 사용한 데이터 기반의 모형 연구 결과로만은 남세균 수를 결정하는 인자들을 명확히 규명하기는 힘들고 추가적 연구가 필요한 것이 사실이나, 보의 영향만으로 기후변화에 대한 취약성을 진단할 수 없음도 분명해 보인다.

상수원 및 친수활동 구간 경계 수준(10,000~100,000 cells/mL)의 남세균 발생은 고도정수처리(Schneider and Bláha, 2020)와 친수활동 접근 제한(U. S. EPA., 2019) 등의 대응이 가능하다. 그럼에도 남세균 수가 경계 수준을 초과하면 고도정수처리 및 친수공간 환경위해성 관리를 위한 추가적인 비용이 발생하므로 이에 대응을 위한 기반 시설 및 제도 정비도 필요해 보인다. 남세균 발생 경계 기준 초과 빈도가 해당 지역의 남세균 발생에 대한 취약도를 나타내므로, 본 연구의 기후변화 시뮬레이션 방법으로 예측한 미래의 남세균 발생 경계 기준 초과 빈도를 낮추는 방향으로 수돗물 남세균 독소처리 시설 등의 사후 대책뿐 아니라 수질 오염 유입 저감 등의 사전 예방 대책을 지역에 맞게 마련하는 것이 가능할 것이다.

#### 4. Conclusion

본 연구에서는 기후변화 시나리오(RCPs)에 따라 2050년과 2100년에 낙동강에서 발생할 남세균의 증식 정도 및 빈도를 기계학습 모형으로 예측했다. 결과적으로 연구 대상 지점들 모두에서 기후변화에 의한 기온상승으로 미래의 남세균 수와 경계 기준 초과 빈도수가 증가하였다. 지금도 남세균과 같은 녹조로 인한 수질 및 수생생태계 문제가 쟁점이 되는 낙동강 지역이 기후변화에 따라 남세균 발생의 정도와 빈도가 증가할 것으로 예측된 것이다. 온실가스 배출이 현재의 추세로 계속 증가하는 경우인 최악의 기후변화 시나리오(RCP8.5)에 의해서는 MM 지점의 경우처럼 남세균 번성의 정도가 “대발생” 수준을 넘어서까지 악화할 수도 있다. 2050년까지 탄소중립 달성하기 위한 신재생에너지 전환을 이루어야 최악의 기후변화 시나리오보다 남세균 수와 경계 초과 빈도를 저감시킬 수 있고 남세균 대발생의 재앙을 예방할 수 있다는 점도 본 연구의 시뮬레이션 결과에서 보였다.

홍수, 가뭄, 해수면 상승, 산사태 등의 자연재해 발생 빈도 및 정도가 기후변화에 따라 증가하는 것처럼, 남세균 증식

현상 빈도가 잦았던 지역이 기후변화에도 취약할 것이라는 가설이 본 연구의 예측 시뮬레이션 결과를 통해서 확인되었다. 또한 경계 기준을 초과하는 남세균 발생 빈도가 기후변화에 따른 남세균 발생 취약성을 대변하고, 이를 관리지표로 사용해서 남세균 발생에 대한 사전 예방과 사후 방지 대책 마련에 기여할 수 있음을 제안하였다.

미래의 기후변화 대응을 위해서는 해당 지역의 남세균 발생 원인에 대한 면밀한 분석을 통해서 남세균 발생을 저감하도록 수질 오염유입 최소화, 유량과 유체 흐름 및 퇴적물 관리 등의 다면적인 대책이 지역에 맞게 마련되어야 한다. 본 연구에서는 기후변화에 따라 예상되는 다양한 기상학적 인자의 변동 중에서도 기온상승 조건만 고려했다는 한계가 있다. 국내에서 녹조 피해가 가장 심한 낙동강 지역의 녹조 대책 마련을 위해서는 남세균 발생 빈도를 줄이기 위한 원인 파악을 위한 지속적인 모니터링 조사연구와 기후변화 시나리오에 따른 기온상승과 더불어 수질 특성, 유량 및 유속, 퇴적도 등 추가적인 인자들이 조합된 복합적인 영향을 시뮬레이션하는 추가 연구가 필요하다.

#### Acknowledgement

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea, funded by the Ministry of Education (No. 2018R1A6A1A08025348).

#### References

- Aggarwal, C. C. (2018). *Neural networks and deep learning*, Springer. <https://doi.org/10.1007/978-3-319-94463-0>
- Bae, D., Shin, S. Y., Shim, J. H., Jeong, S. S., and Shin, S. Y. (2015). The 3rd forum on future policies civil engineering dealing with climate change and urbanization – Preparedness plan for disasters in urban areas, *Journal of the Korean Society of civil engineers*, 63(7), 36-51. [Korean Literature]
- Bernat-Quesada, F., Alvaro, M., Garcia, H., and Navalón, S. (2020). Impact of chlorination and pre-ozonation on disinfection by-products formation from aqueous suspensions of cyanobacteria: *Microcystis aeruginosa*, *Anabaena aequalis* and *Oscillatoria tenuis*, *Water Research*, 183, 116070. <https://doi.org/10.1016/j.watres.2020.116070>
- Cha, Y., Cho, K., Lee, H., Kang, T., and Kim, J. H. (2017). The relative importance of water temperature and residence time in predicting cyanobacteria abundance in regulated rivers, *Water Research*, 124, 11-19. <https://doi.org/10.1016/j.watres.2017.07.040>
- Cho, Y. M., Seo, Y., Maeng, S. K., and Hong, Y. (2020). *Improving the safety of tap water in Gyeonggi-do*, Gyeonggi Research Institute. [Korean Literature]
- Choi, S. Y., Han, K. Y., and Kim, B. H. (2012). Comparison of different multiple linear regression models for real-time flood stage forecasting, *Journal of the Korean Society of*

- Civil Engineers*, 32(1B), 9-20.
- Health Canada. (2022). *Guidelines for Canadian recreational water quality - Cyanobacteria and their Toxins*, Health Canada.
- Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M. H., and Visser, P. M. (2018). Cyanobacterial blooms, *Nature reviews, Microbiology*, 16(8), 471-483. <https://doi.org/10.1038/s41579-018-0040-1>
- Intergovernmental Panel on Climate Change (IPCC). (2014). *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*, Core Writing Team, Pachauri, R. K. and Meyer, L. A. (eds.), Geneva, Switzerland, 151.
- Jöhnk, K. D., Huisman, J., Sharples, J., Sommeijer, B., Visser, P. M., and Stroom, J. M. (2008). Summer heatwaves promote blooms of harmful cyanobacteria, *Global Change Biology*, 14, 495-512. <https://doi.org/10.1111/j.1365-2486.2007.01510.x>
- Jung, W., Kim, S. E., and Kim, Y. D. (2021). Analysis of influential factors of cyanobacteria in the mainstream of Nakdong river using random forest, *Journal of Wetlands Research*, 23(1), 27-34. [Korean Literature] <https://doi.org/10.17663/JWR.2021.23.1.27>
- Jung, Y. J. and Kim, Y. J. (2016). Relationship between energy consumption and operational variables at wastewater treatment plant, *Journal of Korean Society on Water Environment*, 32(3), 253-260. <https://doi.org/10.15681/KSWE.2016.32.3.253>
- Kim, H. G., Cha, Y., and Cho, K. H. (2024). Projected climate change impact on cyanobacterial bloom phenology in temperate rivers based on temperature dependency, *Water Research*, 249, 120928. <https://doi.org/10.1016/j.watres.2023.120928>
- Kim, H. G., Cho, K. H., and Recknagel, F. (2023). Time-series modelling of harmful cyanobacteria blooms by convolutional neural networks and wavelet generated time-frequency images of environmental driving variables, *Water Research*, 246, 120662. <https://doi.org/10.1016/j.watres.2023.120662>
- Kim, J., Jung, W., An, J., Oh, H. J., and Park, J. (2023). Self-optimization of training dataset improves forecasting of cyanobacterial bloom by machine learning, *The Science of the Total Environment*, 866, 161398. <https://doi.org/https://doi.org/10.1016/j.scitotenv.2023.161398>
- Lee, S., Choi, B., Kim, S. J., Kim, J., Kang, D., and Lee, J. (2022). Relationship between freshwater harmful algal blooms and neurodegenerative disease incidence rates in South Korea, *Environmental Health*, 21(1), 1-11. <https://doi.org/10.1186/s12940-022-00935-y>
- Li, Z., Chen, Q., Xu, Q., and Blanckaert, K. (2013). Generalized likelihood uncertainty estimation method in uncertainty analysis of numerical Eutrophication models: Take bloom as an example, *Mathematical Problems in Engineering*, 2013, 701923. <https://doi.org/10.1155/2013/701923>
- Ministry of Environment. (2020). *Korean climate change assessment report 2020*, Publish No. 11-1480000-001691-01, Ministry of Environment. [Korean Literature]
- Ministry of Environment. (2022). *Annual report on algae (green algae) occurrence and response*, Publish No. 11-14800000-001363-10, Ministry of Environment. [Korean Literature]
- Moon, H., Baik, J., Hwang, S., and Choi, M. (2014). Spatial downscaling of grid precipitation using support vector machine regression, *Journal of Korea Water Resources Association*, 47(11), 1095-1105. [Korean Literature]
- National Disaster Management Research Institute. (2023). *Disaster and safety system improvement report: Expansion of heat wave vulnerability management targets due to climate crisis*, Publish No. 11-1741056-000568-01, National Disaster Management Research Institute. [Korean Literature]
- Paerl, H. W. and Huisman, J. (2008). Blooms like it hot, *Science*, 320(5872), 57-58. <https://doi.org/10.1126/science.1155398>
- Schneider, M. and Bláha, L. (2020). Advanced oxidation processes for the removal of cyanobacterial toxins from drinking water, *Environmental Sciences Europe*, 32(1), 94. <https://doi.org/10.1186/s12302-020-00371-0>
- United States Environmental Protection Agency (U. S. EPA.). (2019). *Recommended human health recreational ambient water quality criteria or swimming advisories for microcystins and cylindrospermopsin*, EPA 822-R-19-001. <https://www.epa.gov/sites/default/files/2019-05/documents/hh-rec-criteria-habs-document-2019.pdf>.
- Villanueva, P., Yang, J., Radmer, L., Liang, X., Leung, T., Ikuma, K., Swanner, E. D., Howe, A., and Lee, J. (2023). One-week-ahead prediction of cyanobacterial harmful algal blooms in Iowa lakes, *Environmental Science & Technology*, 57(49), 20636-20646. <https://doi.org/10.1021/acs.est.3c07764>
- Wiśniewska, K., Lewandowska, A. U., and Śliwińska-Wilczewska, S. (2019). The importance of cyanobacteria and microalgae present in aerosols to human health and the environment—Review study, *Environment International*, 131, 104964. <https://doi.org/10.1016/j.envint.2019.104964>
- Woo, C. Y., Yun, S. L., Kim, S. K., and Lee, W. (2020). Occurrence of harmful blue-green algae at algae alert system and water quality forecast system sites in Daegu and Gyeongsangbuk-do between 2012 and 2019, *Journal of Korean Society of Environmental Engineers*, 42(12), 664-673. [Korean Literature] <https://doi.org/10.4491/KSEE.2020.42.12.664>
- World Health Organization (WHO). (2021). *Guidelines on recreational water quality: Volume 1 Coastal and fresh waters*, World Health Organization.
- Xiao, X., He, J., Huang, H., Miller, T. R., Christakos, G., Reichwaldt, E. S., Ghadouani, A., Lin, S., Xu, X., and Shi, J. (2017). A novel single-parameter approach for forecasting algal blooms, *Water Research*, 108, 222-231. <https://doi.org/10.1016/j.watres.2016.10.076>
- Xie, Z., Lou, I., Ung, W. K., and Mok, K. M. (2012). Freshwater algal bloom prediction by support vector machine in Macau storage reservoirs, *Mathematical Problems in Engineering*, 2012, 397473. <https://doi.org/10.1155/2012/397473>