

머신러닝을 이용한 낙동강 본류 구간 수문-기상인자 조류 예보체계 연구

이태우* · 김수전*** · 이준형*** · 김경훈* · 이호용*** · 김덕길****

*인하대학교 토목공학과
**인하대학교 사회인프라공학과
***인하대학교 스마트시티공학전공
****LIG시스템 재난안전연구소

A study on algal bloom forecast system based on hydro-meteorological factors in the mainstream of Nakdong river using machine learning

Taewoo Lee*, Soojun Kim***, Junhyeong Lee***, Kyunghun Kim*, Hoyong Lee***, Duckgil Kim****

**Department of Civil Engineering, Inha University, Korea
***Program in Smart City Engineering, Inha University, Korea
****Disaster and Safety Research Institute, LIG System, Korea

(Received : 28 June 2024, Revised : 09 August 2024, Accepted : 09 August 2024)

요약

남조류의 대량 발생은 담수 생태계에서 수체 내 산소 고갈, 악취 및 독성물질 분비로 인하여 수생태계 및 정수공급체계에 악영향을 미친다. 이러한 녹조현상은 낙동강 보 건설 이후 조류의 수체 내 체류시간 증가와 더불어 기후변화로 인한 지표면 기온 상승으로 인하여 강도와 빈도가 증가할 것으로 예상된다. 본 연구에서는 예상되는 녹조현상의 증가에 대응하기 위하여 조류경보 발령 이전에 선제적 대응을 위한 수문-기상인자 조류 예보체계를 제시하였다. 다연상관 분석을 통하여 조류예측 단계에 따른 기온 및 유량의 선행 영향기간을 탐색하였다. 머신러닝 기법인 의사결정나무 분류를 통하여 선행 기간의 기온 및 유량에 따른 조류예측 단계 분류모델을 도출하였고, 분류모델 결과를 기반으로 수문-기상인자 조류 예보체계를 도출하였다. 제시한 수문-기상인자 조류 예보체계는 녹조현상 발생 이전의 선제적 대응을 위한 기초 연구로서 활용될 수 있을 것으로 판단된다.

핵심용어 : 조류 예보체계, 조류경보제, 남조류, 수문-기상인자, 머신러닝

Abstract

Blue-green algal bloom, or harmful algal bloom has a negative impact on the aquatic ecosystem and purified water supply system due to oxygen depletion in the water body, odor, and secretion of toxic substances in the freshwater ecosystem. This Blue-green algal bloom is expected to increase in intensity and frequency due to the increase in algae's residence time in the water body after the construction of the Nakdong River weir, as well as the increase in surface temperature due to climate change. In this study, in order to respond to the expected increase in green algae phenomenon, an algal bloom forecast system based on hydro-meteorological factors was presented for preemptive response before issuing a algal bloom warning. Through polyserial correlation analysis, the preceding influence periods of temperature and discharge according to the algal bloom forecast level were derived. Using the decision tree classification, a machine learning technique, Classification models for the algal bloom forecast levels based on temperature and discharge of the preceding period were derived. And a algal bloom forecast system based on hydro-meteorological

†To whom correspondence should be addressed.

Department of Civil Engineering, Inha University, Korea
E-mail : sk325@inha.ac.kr

- Taewoo Lee Department of Civil Engineering, Inha University, Korea / Ph.D. Student (dfd0d7cf@nate.com)
- Soojun Kim Department of Civil Engineering, Inha University, Korea / Professor (sk325@inha.ac.kr)
- Junhyeong Lee Program in Smart City Engineering, Inha University, Korea / Ph.D. Candidate (lee_junhyeong.naver.com)
- Kyunghun Kim Department of Civil Engineering, Inha University, Korea / Ph.D. Candidate (tgb611@naver.com)
- Hoyong Lee Program in Smart City Engineering, Inha University, Korea / Ph.D. Student (hy5890@nate.com)
- Duckgil Kim Disaster and Safety Research Institute, LIG System, Korea / Senior Researcher (duckgil.kim@lig.kr)

factors was derived based on the results of the decision tree classification models. The proposed algae forecast system based on hydro-meteorological factors can be used as basic research for preemptive response before blue-green algal blooms.

Key words : Algal bloom forecast system, Algal bloom warning system, Blue-green algae, Hydro-meteorological factor, Machine learning

1. 서론

담수 생태계에서 남조류(Blue-green algae)의 대량 증식 및 사멸에 따라 수체 내 산소 고갈 및 유기물 증가와 같은 문제를 야기하고 있다. 또한 남조류가 분비하는 독성물질 또한 수생태계와 인체에 유해하다. 그리고 인체에는 무해하다고 밝혀졌지만 수돗물 등에서 흙냄새와 같은 좋지 않은 냄새를 유발하는 냄새 물질인 지오스민, 2-MIB을 분비하여 정수 공급체계의 악영향을 미친다.

낙동강 본류 구간에 위치한 8개의 다기능 보 준설로 인하여 하천 수심이 증가하고 유속이 느려지게 되었다. 국내외 연구자는 저수지나 보가 건설되는 경우 하천의 유속 저하로 인한 수리학적 체류시간 증가는 조류의 성장에 유리한 환경이 될 수 있다는 것을 보고한 바 있다(Gore and Pettis, 1989; Joo and Jeong, 2005; Kim et al., 2018). 2013년부터 운영하는 조류경보제에 의하여 운영되는 낙동강 본류 구간의 조류경보 지점 4개소(해평, 강정고령, 칠서, 물금-매리)는 매년 주의보 또는 경보가 발령되고 있다. 이러한 녹조 문제가 대두됨에 따라 하천 및 호소에서의 남조류 성장 원인을 규명하기 위하여 인공지능 또는 머신러닝 기법을 활용한 다수의 연구가 수행되었다. Park et al.(2021)은 인공지능경망과 Support Vector Machine(SVM) 기법을 통하여 저수지의 조류 발생 예측을 수행하였으며, 수질 및 기상, 수문학적 인자의 영향을 많이 받으며 조기 예보 일자로 6일에서 7일을 제안하였다. Heddam et. al.(2022)은 미국에 위치한 하천에서의 남조류를 모델링하기 위하여 인공지능경망, Extreme Learning Machine, 랜덤 포레스트 회귀(Random Forest Regressor, RFR), Random Vector Function Link 기법을 적용하였으며 랜덤 포레스트 모델의 우수한 예측 정확도를 확보하였다. Nguyen et. al.(2020)은 위성자료와 가우시안 프로세스 회귀(Gaussian Process Regressor, GPR), RFR, SVM, Multi-Layer Perceptron Regressor (MLP) 등의 머신러닝 기법을 통하여 베트남 남부에 위치한 저수지에서의 Chl-a 농도 예측 모델을 제시하였으며, Chl-a 농도 증가의 주된 요인으로 장마철 풍부한 강수량으로 인한 영양염류 유입을 지목하였다. Ly et. al.(2021)은 한강 본류의 부영양화 예측을 위하여 Adaptive Neuro-Fuzzy Inference System(ANFIS) 기법을 통하여 수질인자 및 기온, 강수량, 유량 인자로부터 Chl-a 농도를 예측하였다. Jeong et. al.(2022)은 RFR과 Extreme Gradient Boosting 기법을 통하여 국내 8개 저수지의 유해 남조류를 예측하였으며 수온과 총질소, 총인 농도가 유해 남조류 발생에 큰

영향을 미치는 것을 확인하였다.

이러한 녹조 문제는 미래 기후변화로 인한 지표면 온도 상승으로 인하여 더욱 심화될 것으로 전망된다. Intergovernmental Panel on Climate Change(IPCC)에 따르면 지구 표면의 온도는 인간 활동으로 인한 온실가스 배출로 인하여 1850년~1900년 이후 2021년까지 약 1.1°C 상승하였으며, Shared Socioeconomic Pathway(SSP) 시나리오 중 SSP1-1.9 및 SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5 시나리오에서 향후 20년 동안 가장 유력한 평균 지구 온도 상승량을 1.5°C~1.6°C로 전망하였다(IPCC, 2021). 이러한 지구 온도 상승은 유역 내 호수 및 저수지의 유해 조류 확산(Harmful algal blooms)으로 이어지게 된다고 주장되고 있다(Paerl and Huisman, 2009; Visser et al., 2016; Huisman et al., 2018). 기후변화로 인한 기온 상승은 하천 및 호소 내 서식하는 조류 중 유해 조류인 남조류(Cyanobacteria)의 우점을 초래하고(Paerl and Huisman, 2008; Carey et al., 2012; Visser et al., 2016), 여름철 식생 성장 기간의 연장으로 이어지며(Anneville et al. 2005; Deng et al. 2014; Visser et al. 2016), 호수의 성층화(Lake stratification)를 유발하여 부력 조절이 가능한 남조류의 생존성을 높게 된다(Wagner and Adrian, 2009; Carey et al., 2012; Posch et al., 2012; Taranu et al., 2012; Visser et al., 2016).

이렇듯 기후변화로 인해 녹조현상의 빈도와 강도가 극심해질 것으로 예상됨에 따라 사전 녹조 예측을 통한 선제적인 관리가 필요한 실정이다. 이에 본 연구에서는 남조류 성장에 주된 영향을 미치는 수문인자인 유량과 기상인자인 기온 요소로부터 영향을 미치는 선행 기간을 탐색하고, 선행 기간을 고려한 기온과 유량 인자를 기준으로 머신러닝 기법 중 하나인 의사결정나무 분류분석을 수행하여 분류모델을 도출한다. 도출한 분류모델로부터 조류예측 단계별 기온·유량 임계 조건을 추정하고, 추정된 기온·유량 임계조건을 기반으로 수문-기상 조류 예보체계를 제시한다.

2. 연구방법

2.1 조류경보제(Algal bloom warning system) 및 조류 예측(Algal bloom forecast)

조류경보제는 조류 발생 상황을 주기적으로 모니터링하여 일정 수준 이상의 조류가 발생할 경우 경보를 발령하여 필요한 조치를 취함으로써 수돗물을 안정적으로 공급하고, 친수활동시 조류독소로부터 국민의 안전을 도모하기 위해 물환경보전법 제21조에 근거하여 시행되고 있다. 또한 사

Table 1. Standards according to the algal bloom forecast level in South Korea

Algal bloom level	Cyanobacteria cell count standards according to level
0	less than 1,000 cell/ml
1	1,000 cell/ml or more but less than 10,000 cell/ml
2	10,000 cell/ml or more but less than 100,000 cell/ml
3	100,000 cell/ml or more but less than 1,000,000 cell/ml
4	1,000,000 cell/ml or more

Table 2. Algae, climate and water level observation station in this study

No.	Algal bloom warning station	Climate station		Water level station	
		Name	Type	Name	Type
1	Haepyeong (HP)	Gumi	ASOS	Gumi weir	Reservoir
2	Gangjung Goryeong (GG)	Daegu	ASOS	Seongju bridge	River
3	Chilseo (CS)	Uiryong	ASOS	Gyenaeri	River

전 예방적 수질관리와 조류 발생 대응 등에 활용할 수 있도록 동법 시행령 제28조에 근거하여 조류정보를 예측하기 위하여 조류발생예측시스템이 운영되고 있다. 조류발생예측시스템에서 준용하는 조류예측 단계는 조류예측 및 수질관리협의회에 관한 규정 제8조 제3항에서 제시하고 있다. 유역의 호소 및 하천 내 남조류(Cyanobacteria)의 단위 체적당 세포수를 기준으로 0단계부터 4단계까지 총 5가지 단계로 구분하며 구체적인 세부 기준은 Table 1에 나타내었다. 남조류세포수는 마이크로시스티스(Microcystis), 아나베나(Anabaena), 아파니조메논(Aphanizomenon), 오실라토리아(Oscillatoria) 속들의 세포수 합을 기준으로 한다. 조류예측 발표 기간은 조류예측 및 수질관리협의회에 관한 규정 제8조 제1항에 따라 기온이 높아 조류 생장에 유리한 기간인 매년 5월부터 10월까지 시행한다. 시행지점은 전국 총 29개소로 상수원 호소 22개소, 상수원 하천 6개소, 친수활동구간 1개소로 구성되어 있으며, 낙동강에서는 상수원 호소 9개, 낙동강 본류 하천 4개 구간에서 운영하고 있다.

2.2 대상지역 및 사용 데이터

본 연구에서는 녹조 발생 빈도가 높은 낙동강 유역을 대상유역으로 채택하였으며, 낙동강 본류 하천구간 4개소 중 관측자료 보유기간이 5년 이상인 해평(Haepyeong, HP) 지점과 강정고령(Gangjung Goryeong, GG) 지점, 칠서(Chilseo, CS) 지점을 선정하였다. 본 연구에서는 물환경정보시스템(<http://water.nier.go.kr>)에서 제공하는 물환경측정망 자료 중 자료제공 시작년도인 2016년 상반기부터 2022년까지, 조류예측 발표가 시행되는 5월부터 10월까지의 기간의 관측자료를 적용하였다. 또한 남조류 생장에 영향을 미치는 수문-기상인자 분석을 위해 조류정보지점에서 가장 인접한 기상청 종관기상관측소(ASOS) 및 수위·유량 관측소에서 제공하는 일평균기온(Temperature), 유량(Discharge) 시계열 자료를 활용하였다. 본 연구에서 선정한 조류정보제 지점 및 기상, 수위·유량 관측소의 위치는 Fig. 1과 같으며,

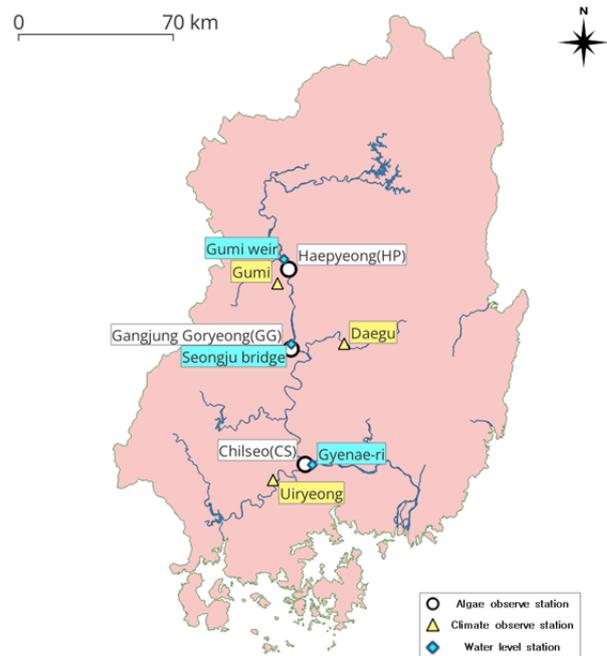


Fig. 1 Study area and algae observe station

조류정보제 지점과 각각의 기상, 수위·유량 관측소 목록은 Table 2에 정리하였다.

2.3 의사결정나무(Decision tree)

의사결정나무는 의사결정규칙(decision rule)을 도표화하여 관심 대상이 되는 집단을 몇 개의 소집단으로 분류(classification)하거나 예측(prediction)을 수행하는 분석방법이다(Jung et. al., 2019). 일반적으로 의사결정나무 분석은 다음과 같은 단계를 거친다. 첫 번째, 의사결정나무의 형성을 통해 분석의 목적과 자료구조에 따라서 적절한 분리 기준(split criterion)과 정지규칙(stopping rule)을 지정하여 의사결정나무를 얻는다. 두 번째, 가지치기를 통해 분류오류(classification error)를 크게 할 위험이 높거나 부적

절한 규칙을 가지고 있는 가치를 제거한다. 세 번째, 타당성 평가를 통해 이익도표나 위험도표 또는 검정용 자료에 의한 교차타당성(cross validation) 등을 이용하여 의사결정나무를 평가한다. 마지막으로, 의사결정나무를 해석하고 분류 및 예측모형을 설정한다(Berry and Linoff, 1997). 본 연구에서는 범주형 변수의 분류를 위하여 의사결정나무의 Classification Trees(CT) 알고리즘을 사용하였다. 조류예측 단계를 범주형 목표변수로 설정하고 주요 영향인자에 따른 남조류 발생 조건을 분석하였다.

3. 의사결정나무 분류모델 기반의 조류 예보체계 제시

3.1 수문-기상인자의 선행 영향기간 탐색을 위한 상관관계 분석

본 연구에서는 주 단위의 시간 간격으로 측정하는 남조류 세포수에 대하여 기온 및 유량이 남조류의 성장에 영향을 미치는 선행 기간을 고려하기 위하여 상관관계 분석을 수행하였다. 최적의 선행기간을 탐색하기 위하여 관측일 당일의 관측값 및 관측일로부터 1주일, 2주일, 3주일, 4주일 전까지의 기온과 유량 관측값의 평균값을 산정하여 각각의 인자와 조류예측 단계 간의 상관관계 분석을 수행하였다. 남조류 세포수의 특성상 발생 범위가 0 cell/ml부터 1,000,000 cell/ml 또는 그 이상까지 굉장히 넓기 때문에 해당 인자를 연속형 변수로 취급하여 pearson correlation 분석 시 상관관계를 온전히 나타내지 못할 우려가 있다. 따라서 남조류 세포수를 Table 1에 제시한 기준에 따른 조류예측 단계로 치환하여 상관관계 분석을 수행하였다. 범주형 변수인 조류예측 단계와 연속형 변수인 기온, 유량의 상관관계 분석을 수행하기 위하여 다연상관(Polyserial correlation)분석을 수행하였다. 다연상관분석은 범주형 변수와 연속형 변수 간의 상관관계를 규명할 수 있는 상관분석 방법이다(Dragsow, 2004).

대상 관측소의 선행 기간별 기온, 유량과 조류예측 단계 간 다연상관분석을 수행한 결과는 Table 3 및 Fig. 2와 같다. 기온(Temperature) 인자의 경우 전체 관측소에서 +0.45부터 +0.64까지 양의 상관계수가 산정되었으며, 유량(Discharge) 인자의 경우 전체 관측소에서 -0.02부터 -0.42까지 음의 상관계수가 산정되었다. 즉, 기온이 높고 유량이 낮을수록, 즉 수역이 정체될수록 남조류 세포수가 증가할 확률이 높

다는 선행연구의 결과를 따르는 것으로 판단된다. HP 지점의 유량 상관계수가 매우 낮게 산정되었는데, 이는 해당 지점의 유량 자료가 보(Weir)에 의해 통제되는 저수지(Reservoir)의 방류량 자료이기 때문인 것으로 사료된다. 보에 의해 직접 유량이 조절되지 않는 하천(River) 유량 자료를 활용한 GG 지점과 CS 지점의 경우 HP 지점 대비 상대적으로 유량에 대한 상관계수가 높게 산정되었다. 선행기간 관점에서 보았을 때, 기온의 경우 GG 지점을 제외하고는 3주 선행기간까지 상관계수가 증가하였다가 4주 선행기간에서 감소하였고, 유량의 경우 HP 지점을 제외하고는 2주 선행기간까지 상관계수가 증가하였다가 3주 선행기간부터 감소하였다. 이에 본 연구에서는 낙동강 본류 조류경보 지점의 남조류 세포수에 영향을 미치는 기온의 선행 영향기간은 3주, 유량의 선행 영향기간은 2주로 설정하였다.

3.2 의사결정나무를 이용한 남조류 발생 특성 분석

본 연구에서는 기온과 유량에 따른 남조류 발생 조건을 살펴보기 위해 기온과 유량을 독립변수, 남조류 세포수에 따른 조류예측 단계를 종속변수로 설정한 후 의사결정나무를

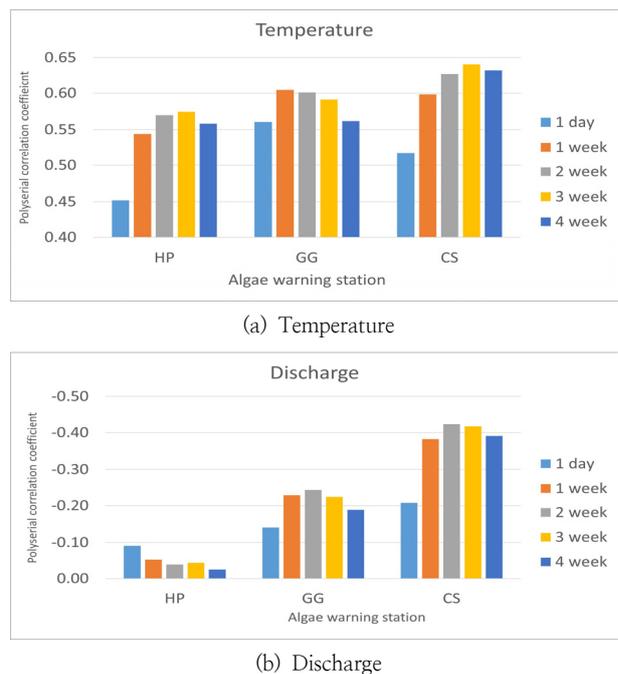


Fig. 2 Visualization of polyserial correlation coefficient between the algal bloom warning level and hydro-meteorological factor(temperature, discharge)

Table 3. Polyserial correlation coefficient between the algal bloom warning level and hydro-meteorological factor(temperature, discharge)

No.	Algae warning station	Polyserial correlation coefficient by preceding period									
		Temperature					Discharge				
		1 day	1 week	2 week	3 week	4 week	1 day	1 week	2 week	3 week	4 week
1	HP	0.45	0.54	0.57	0.57	0.56	-0.09	-0.05	-0.04	-0.04	-0.02
2	GG	0.56	0.60	0.60	0.59	0.56	-0.14	-0.23	-0.24	-0.22	-0.19
3	CS	0.52	0.60	0.63	0.64	0.63	-0.21	-0.38	-0.42	-0.42	-0.39

Table 4. The criteria value and frequency(3 or more time) of temperature and discharge factor from algal bloom warning level classification results

No.	Algae observation station	Temperature (°C)		Discharge (m³/s)	
		Criteria value	Frequency	Criteria value	Frequency
1	HP	20.8	3	144.2	3
		21.1	3	225.3	4
		21.4	4	-	
		22.3	3		
		22.8	4		
		26.2	3		
2	GG	19.9	3	58.0	3
		21.9	6	59.6	3
		24.1	3	63.0	3
		26.9	4	72.3	3
		27.3	3	87.3	3
		28.6	3	92.2	4
		30.4	3	151.7	3
		-		160.4	3
		-		260.7	3
		-		409.7	3
3	CS	14.6	5	110.8	3
		20.4	4	144.3	3
		20.5	4	144.5	4
		20.7	4	176.3	3
		25.5	5	184.1	4
		29.6	5	355.5	3
		-		539.5	3
		-		1030.9	3
		-		1123.3	3
		-		2945.7	3

구축하고 조류예측 단계별 발생 조건을 분석하였다. 의사결정나무 머신러닝을 위한 학습(Training) 세트 자료는 전체 자료 중 70%, 검증(Validation) 세트는 30%로 구성하였으며, 학습 자료 세트와 검증 자료 세트를 무작위로 교체하면서 총 10회 의사결정나무 분류모델을 도출하였다. Fig. 3은 각각의 지점별 총 10회 도출한 의사결정나무 분류모델 중 하나씩을 예시로 나타내었으며, x1 변수는 기온, x2 변수는 유량을 의미하고 0, 1, 2는 조류예측 단계를 의미한다.

도출된 의사결정나무 분류모델을 통해 도출된 모든 뿌리마디 및 내부 마디의 기온과 유량 분리 기준을 취합하였다. 의사결정나무 반복 분석을 통해 특정 분리 기준값이 자주 도출될수록 해당 기준값이 조류예측 단계를 판가름하는 기준으로서 유의미하다고 판단하였으며, 이에 10회 반복 분석 동안 도출 빈도 3회 이상의 기온 및 유량 분리 기준값을 Table 4에 제시하였다.

3.3 의사결정나무 결과 기반 수문-기상인자 조류 예보체계 도출 및 평가

도출한 의사결정나무 분류모델은 변수별로 세밀하고 좁은 범위의 분류기준으로 인하여 조류 예보체계로써 직접

활용하기에는 사용 과정이 매우 복잡하다. 이에 본 연구에서는 보다 일반적인 기준으로 간단히 적용할 수 있는 조류 예보체계를 제시하였다. 이전 절에서 도출한 의사결정나무 분류모델을 통하여 최종 도출한 조류예측 단계에 따른 기온·유량 인자별 임계값을 통하여 HP, GG, CS 조류경보 지점별 수문-기상인자 조류 예보체계를 도출하였다(Fig. 4). 조류예측 단계를 기준으로 Normal 단계는 0단계, Attention 단계는 1단계, Warning 단계는 2단계 이상을 의미한다. 각각의 단계별 구분 기준 온도 및 유량은 Table 4에 도출된 고빈도 분리 기준값으로부터 추출하였으며, 2016년부터 2022년까지 발령된 조류경보 기록을 기준으로 가장 정확도가 높은 기준값을 선별하여 적용하였다. 수문-기상인자 조류 예보체계의 조류예측 단계 평가방법은 다음과 같다. Fig. 4 (a) HP 관측소를 예시로, 구미(Gumi) ASOS 관측소의 지난 3주간의 평균기온이 21.4°C 이상 26.2°C 미만이고 구미보(Gumi weir)의 지난 2주간의 평균 유량이 144.2m³/s 이상 225.3m³/s 미만일 경우 조류 예보체계로부터 예측되는 조류예측 단계는 Attention 등급이다. 이러한 간단한 방식으로 수문-기상인자 조류 예보체계로부터 3주간의 일 평균기온의 평균값과 2주간의 평균 유량으로부터 조류예측 단계를 예측할 수 있다.

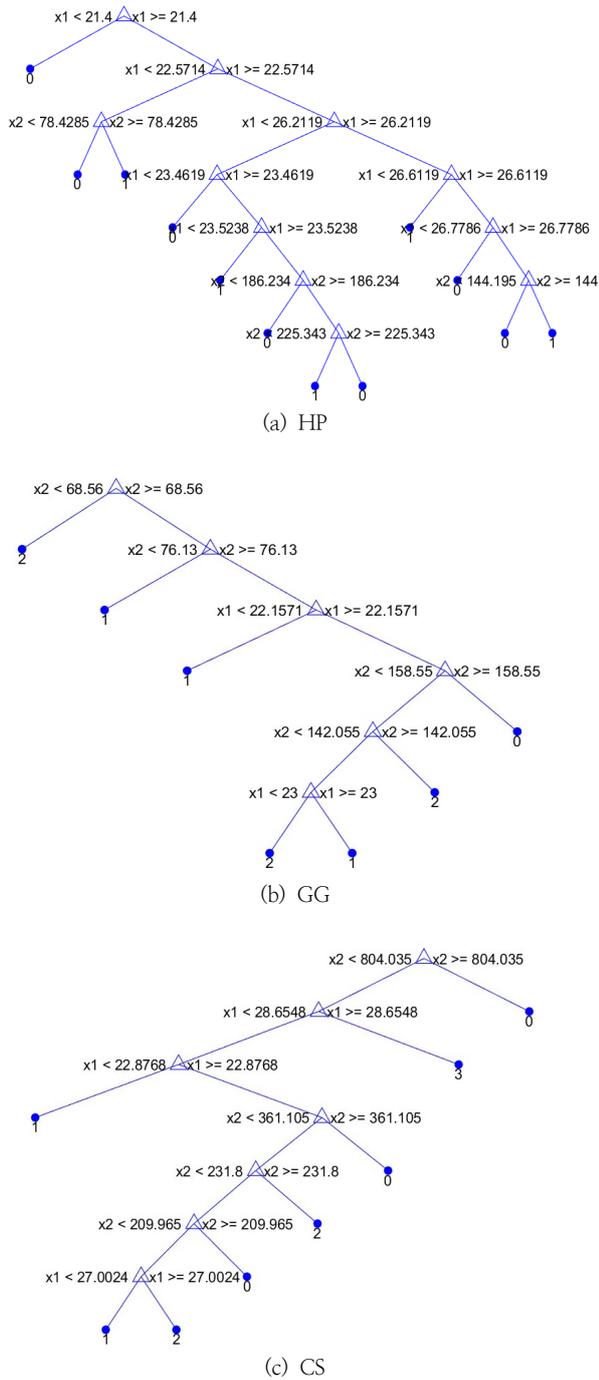


Fig. 3 Example of decision tree result between algal bloom warning level and hydro-meteorological factor

2016년부터 2022년의 5월부터 10월까지를 학습 기간 (Training period), 2023년 5월부터 10월까지를 검증 기간 (Verification period)으로 설정하여 본 연구에서 제시한 조류 예보체계를 검증한 결과는 Fig. 5와 같다. 검증 정확도는 실제 관측된 조류 발생 단계 대비 예측한 조류 발생 단계가 일치하는 건수를 비율로 산정하였다. Fig. 5 (a) HP 지점을 예를 들어, 실제 Normal 단계로 관측된 건(총 143건) 중 조류 예보체계를 통하여 예측된 Normal 단계가 99건, Attention 단계가 37건, Warning 단계가 7건임을 의미

하며, 실제 Normal 관측 건수인 143건 대비 조류 예보체계를 통하여 Normal로 예측된 건수 99건의 비율인 69%를 Normal 단계 예측 정확도로 산정한다. 최종적으로 각각의 모든 단계를 정확히 예측한 건수의 비율을 통합 정확도로 지정하였으며 HP 지점의 학습 기간에서의 통합 정확도는 63.6%로 산정된다. 모든 지점에 대한 정확도를 정리하면 학습기간의 정확도는 63.6%~66.4%, 검증기간의 정확도는 53.6%~76.9%로 산정되었다. 학습기간(Fig. 5(a), (b), (c))과 검증기간(Fig. 5(d), (e), (f))의 단계별 정확도를 비교하면, 두 기간 모두 Normal 단계의 예측 정확도는 비교적 준수한 것으로 판단된다. 그러나 상대적으로 중요한 Attention 단계의 경우 HP 지점 0%, GG 지점 28.6%, CS 지점 18.2%로 산정되고 Warning 단계의 경우 GG 지점 및 CS 지점 모두 0%로 정확도가 다소 낮게 산정되었다. 이에 본 연구에서 제시한 수문-기상인자 조류 예보체계는 실무에 적용하기에는 한계가 있으며 보완이 필요한 것으로 판단된다. 특히 실제 Attention 단계를 Normal 단계로, 실제 Warning 단계를 Attention 단계로 과소추정하는 경향을 보완하는 방안이 필요하다.

HP algae observation station		Gumi ASOS Average temperature for the preceding 3 weeks (°C)		
		21.4	26.2	
Gumi weir Average discharge for the preceding 2 weeks (m³/s)	Normal	Normal	Attention	Warning
		Warning	Normal	Attention
	Warning	Normal	Attention	Warning

(a) HP

GG algae observation station		Daegu ASOS Average temperature for the preceding 3 weeks (°C)		
		21.9	26.9	
Seongju bridge Average discharge for the preceding 2 weeks (m³/s)	Normal	Normal	Attention	Warning
		Warning	Normal	Attention
	Warning	Normal	Attention	Warning

(b) GG

CS algae observation station		Uiryeong ASOS Average temperature for the preceding 3 weeks (°C)		
		20.5	25.5	
Gyenaeri Average discharge for the preceding 2 weeks (m³/s)	Normal	Normal	Attention	Warning
		Warning	Normal	Attention
	Warning	Normal	Attention	Warning

(c) CS

Fig. 4 Hydro-meteorological algal bloom forecast system of algae observation station

HP Training result		Forecasted			Forecast accuracy	
		Normal	Attention	Warning	by level	overall
Observed (187)	Normal (143)	99	37	7	69%	63.6%
	Attention (35)	16	12	7	34%	
	Warning (9)	0	1	8	89%	

(a) HP station in training period

HP Verification result		Forecasted			Forecast accuracy	
		Normal	Attention	Warning	by level	overall
Observed (26)	Normal (22)	19	3	0	86.4%	73.1%
	Attention (4)	4	0	0	0%	
	Warning (0)	0	0	0	- (100%)	

(d) HP station in verification period

GG Training result		Forecasted			Forecast accuracy	
		Normal	Attention	Warning	by level	overall
Observed (191)	Normal (95)	79	15	1	83%	64%
	Attention (66)	22	28	11	46%	
	Warning (30)	0	18	12	40%	

(b) GG station in training period

GG Verification result		Forecasted			Forecast accuracy	
		Normal	Attention	Warning	by level	overall
Observed (26)	Normal (18)	18	0	0	100%	76.9%
	Attention (7)	5	2	0	28.6%	
	Warning (1)	0	1	0	0%	

(e) GG station in verification period

CS Training result		Forecasted			Forecast accuracy	
		Normal	Attention	Warning	by level	overall
Observed (223)	Normal (80)	76	4	0	95%	66.4%
	Attention (67)	26	37	4	55%	
	Warning (76)	8	33	35	46%	

(c) CS station in training period

CS Verification result		Forecasted			Forecast accuracy	
		Normal	Attention	Warning	by level	overall
Observed (28)	Normal (13)	13	0	0	100%	53.6%
	Attention (11)	9	2	0	18.2%	
	Warning (4)	0	4	0	0%	

(f) CS station in verification period

Fig. 5 Forecast accuracy of hydro-meteorological algal bloom forecast system in training period and verification period

4. 결론

본 연구의 대상지역인 낙동강 유역은 다기능 보 건설 이후 수심 증가 및 유속 저하로 인하여 보 건설 이전보다 하천 및 호소의 남조류 생장이 유리한 환경을 갖게 되었다. 또한 기후변화로 인하여 예상되는 지표면 기온 증가는 향후 하천 및 호소에 녹조현상의 강도 및 빈도 증가에 직접적인 영향을 미칠 것으로 예상된다. 이에 본 연구에서는 낙동강 본류의 조류경보 지점에 대하여 남조류 생장에 영향을 미치는 수문-기상인자인 기온과 유량에 대한 선행 영향기간을 탐색하고, 머신러닝 기법인 의사결정나무 분류모델을 구축하여 구축한 분류모델을 기반으로 조류예측 단계에 따른 기온과 유량 구분 기준을 도출하였다. 최종적으로 조류예측 단계별 기온 및 유량 구분 기준을 통하여 간단하게 조류예측 단계를 예측할 수 있는 조류 예보체계를 제시하였다. 본 연구의 결과를 요약하면 다음과 같다.

(1) 수문-기상인자가 남조류 생장에 영향을 미치는 선행기간을 탐색하기 위하여 낙동강 본류의 대상 조류경보 지점의 인근 기상 관측소의 기온 자료 및 수위·유량 관측소의 유량 자료를 토대로 조류예측 단계와의 상관관계 분석을 수행하였다. 연속형 자료인 기온 및 유량과 범주형 자료인 조류예측 단계와의 상관관계 분석을 위해서 다연상관(Polyserial correlation) 분석을 수행하였다. 관측일 당일 및 1주, 2주, 3주, 4주의 선행 기간별 평균값으로 상관관계

분석을 수행한 결과 기온의 경우 3주, 유량의 경우 2주의 선행 기간이 상관관계가 가장 높게 산정되었다. 이에 해당 선행 기간부터 조류 발생 예보가 수행된다면 녹조현상 예방에 기여할 수 있을 것으로 판단된다.

(2) 기온 및 유량의 선행 영향기간 평균값과 조류 측정 자료로부터 의사결정나무 분류모델을 도출하였다. 총 10회 도출한 의사결정나무 분류모델로부터 조류예측 단계에 따른 기온과 유량의 모든 분리 기준값을 추출하였다. 도출한 분리 기준값들을 기반으로 수문-기상 조류 예보체계를 제시하였다. 모든 조류경보 지점에 대한 수문-기상 조류 예보체계의 전체 단계에 대한 실제 발령단계와 모의 발령단계가 일치하는 지에 대한 정확도를 평가한 결과, 학습기간에서 63.6%~66.4%, 검증기간에서 53.6%~76.9%로 산정되었다. 그러나 검증기간의 경우, Normal 단계를 제외한 Attention 단계와 Warning 단계의 정확도가 0%~28.6% 수준으로 매우 저조함에 따라 보완이 필요한 것으로 판단된다.

본 연구에서 제시한 수문-기상인자 조류 예보체계를 조류경보 지점별로 기온과 유량의 수치적 기준에 따른 조류예측 단계를 제시함으로써 지점에 따른 맞춤형 남조류 관리 위한 선제적 의사결정을 위한 기반 연구로 활용할 수 있을 것으로 판단된다. 다만, 고위험 단계에 대한 낮은 예측 정확도로 인하여 정확도 제고를 위한 보완이 필요하다. 남조류세포수의 측정주기는 1주일로, 관측 시작년도인 2016년을 기준으로 2023년까지의 조류 경보기간인 5월부터

터 10월간 시계열 자료의 개수가 약 208여 개만이 존재하므로 머신러닝 학습을 위한 자료량이 다소 부족한 것으로 판단된다. 또한 남조류의 생장 조건은 기상-수문-수질 요인이 복합적으로 작용하기 때문에 수질인자를 고려할 필요가 있으나, 수질인자의 경우 다음과 같은 문제점이 존재한다. 우선, 조류 경보지점에서는 총질소, 총인 등 일부 수질인자를 측정하지 않기 때문에 수질인자 자료를 수질측정망과 자동측정망 두 측정소 유형으로부터 확보할 수 있다. 수질측정망의 경우, 측정소 위치의 밀도가 높지만 측정주기가 주 단위이기 때문에 자료계열 간 연속성이 보존되지 않으며 측정 이후 확정자료가 나오기까지 약 2개월이 소요되기 때문에 예보체계 활용에 부적합한 것으로 판단된다. 자동측정망의 경우 매시간 측정하여 자료계열의 연속성을 확보할 수 있으나, 최근 2개월까지는 검증 없이 측정된 수치 그대로 미확정 자료를 제공하며 확정자료를 제공하기까지 약 2개월이 소요된다. 그리고 자동측정망의 측정소 위치의 밀도가 다소 낮아 각각의 자동측정망 측정소의 위치가 조류 경보지점과 멀리 떨어져 있기 때문에 해당 수질자료를 그대로 활용하기 어려울 것으로 판단된다. 향후 연구에서는 수질측정망 확정자료를 활용하여 기상-수문-수질 요인이 복합적으로 고려된 조류 예보체계를 제시하고, 정확도를 유의미한 수준으로 개선하여 예보 활용에 대한 타당성이 입증될 경우 미확정 자료 활용을 통하여 현행 조류 예경보에 활용될 수 있을 것으로 기대된다.

감사의 글

본 결과물은 환경부의 재원으로 한국환경산업기술원의 “기후위기대응 홍수방어능력 기술개발사업”의 지원을 받아 연구되었습니다.(2022003460002)

References

- Anneville, O., Gammeter, S., and Straile, D. (2005). “Phosphorus decrease and climate variability: mediators of synchrony in phytoplankton changes among European peri-alpine lakes.” *Freshwater Biology*, Vol. 50, No. 10, pp. 1731–1746.
- Berry, M. J. A., and Linoff, G. S. (1997). *Data Mining Techniques*, New York, John Wiley & Sons Publisher
- Carey, C. C., Ibelings, B. W., Hoffmann, E. P., Hamilton, D. P., and Brookes, J. D. (2012). “Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate.” *Water research*, Vol. 46, No. 5, pp. 1394–1407.
- Deng, J., Qin, B., Paerl, H. W., Zhang, Y., Ma, J., and Chen, Y. (2014). “Earlier and warmer springs increase cyanobacterial (*Microcystis* spp.) blooms in subtropical Lake Taihu, China.” *Freshwater Biology*, Vol. 59, No. 5, pp. 1076–1085.
- Drasgow, F. (2004). *Polychoric and polyserial correlations*. Encyclopedia of statistical sciences.
- Gore, J. A., and Pettis, G. E. (1989). *Alternatives in regulated river management*. CRC Press, F.L., U.S.
- Heddam, S., Yaseen, Z. M., Falah, M. W., Goliatt, L., Tan, M. L., Sa’adi, Z., Ahmadianfar, I., Saggi, M., Bhatia, A., and Samui, P. (2022). “Cyanobacteria blue-green algae prediction enhancement using hybrid machine learning-based gamma test variable selection and empirical wavelet transform.” *Environmental Science and Pollution Research*, Vol. 29, No. 51, pp. 77157–77187.
- Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M., and Visser, P. M. (2018). “Cyanobacterial blooms.” *Nature Reviews Microbiology*, Vol. 16, No. 8, pp. 471–483.
- Intergovernmental Panel on Climate Change (IPCC). (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. In V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Eds.), *2021: Summary for Policymakers*. (pp. 3–32) Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/9781009157896.001.
- Jeong, B., Chapeta, M. R., Kim, M., Kim, J., Shin, J., & Cha, Y. (2022). “Machine learning-based prediction of harmful algal blooms in water supply reservoirs.” *Water Quality Research Journal*, Vol. 57, No. 4, pp. 304–318.
- Joo, G., and Jeong, K. (2005). *Modelling community changes of cyanobacteria in a flow regulated river (the lower Nakdong River, S. Korea) by means of a Self-Organizing Map (SOM)*. Modelling Community Structure in Freshwater Ecosystems. Springer, Berlin, Germany, pp. 273–287.
- Jung, W. S., Kim, B. G., Kim, Y. D., and Kim, S. E. (2019). “A study on the characteristics of cyanobacteria in the mainstream of Nakdong river using decision trees.” *Journal of Wetlands Research*, Vol. 21, No. 4, pp. 312–320.
- Kim, J.S., Seo, I.W., and Baek, D. (2018). “Modeling spatial variability of harmful algal bloom in regulated rivers using a depthaveraged 2D numerical model.” *Journal of Hydro-environment Research*, Vol. 20, pp. 63–76.
- Li, H. M., Tang, H. J., Shi, X. Y., Zhang, C. S., and Wang, X. L. (2014). “Increased nutrient loads from the Changjiang (Yangtze) River have led to increased harmful algal blooms”. *Harmful Algae*, Vol. 39, pp. 92–101.
- Ly, Q. V., Nguyen, X. C., Le, N. C., Truong, T. D., Hoang, T. H. T., Park, T. J., Maqbool, T., Pyo, J. C., Cho, H. K., Lee,

- K. S., and Hur, J. (2021). "Application of Machine Learning for eutrophication analysis and algal bloom prediction in an urban river: A 10-year study of the Han River, South Korea." *Science of The Total Environment*, Vol. 797, pp. 149040.
- Nguyen, H. Q., Ha, N. T., & Pham, T. L. (2020). "Inland harmful cyanobacterial bloom prediction in the eutrophic Tri An Reservoir using satellite band ratio and machine learning approaches." *Environmental Science and Pollution Research*, Vol. 27, pp. 9135–9151.
- Paerl, H. W., and Huisman, J. (2008). "Blooms like it hot." *Science*, Vol. 320, Issue 5872, pp. 57–58.
- Paerl, H. W., and Huisman, J. (2009). "Climate change: a catalyst for global expansion of harmful cyanobacterial blooms." *Environmental microbiology reports*, Vol. 1, No. 1, pp. 27–37.
- Park, Y., Lee, H. K., Shin, J. K., Chon, K., Kim, S., Cho, K. H., Kim, J. H., and Baek, S. S. (2021). "A machine learning approach for early warning of cyanobacterial bloom outbreaks in a freshwater reservoir." *Journal of Environmental Management*, Vol. 288, pp. 112415.
- Posch, T., Koster, O., Salcher, M. M., and Pernthaler, J. (2012). "Harmful filamentous cyanobacteria favoured by reduced water turnover with lake warming." *Nature Climate Change*, Vol. 2, No. 11, pp. 809–813.
- Taranu, Z. E., Zurawell, R. W., Pick, F., and Gregory-Eaves, I. (2012). "Predicting cyanobacterial dynamics in the face of global change: the importance of scale and environmental context." *Global Change Biology*, Vol. 18, No. 12, pp. 3477–3490.
- Visser, P. M., Verspagen, J. M., Sandrini, G., Stal, L. J., Matthijs, H. C., Davis, T. W., Paerl, H. W., and Huisman, J. (2016). "How rising CO₂ and global warming may stimulate harmful cyanobacterial blooms." *Harmful algae*, Vol. 54, pp. 145–159.
- Wagner, C., and Adrian, R. (2009). "Cyanobacteria dominance: quantifying the effects of climate change." *Limnology and Oceanography*, Vol. 54, Issue 6 part2, pp. 2460–2468.