

다중회귀분석을 이용한 낙동강 하류의 Chlorophyll-a 농도와 복합 영향인자들의 상관관계 분석

Evaluation of Correlation between Chlorophyll-a and Multiple Parameters by Multiple Linear Regression Analysis

임지성 · 김영우 · 이재호 · 박태주 · 변임규*[†]

Ji-Sung Lim · Young-Woo Kim · Jae-Ho Lee · Tae-Joo Park · Im-Gyu Byun*[†]

부산대학교 사회환경시스템공학부 · *부산대학교 환경기술·산업개발연구소
Department of Civil and Environmental Engineering, Pusan National University
Institute for Environmental Technology and Industry, Pusan National University

(Received March 23, 2015; Revised April 23, 2015; Accepted May 11, 2015)

Abstract : In this study, Chlorophyll-a (chl-a) prediction model and multiple parameters affecting algae occurrence in Mulgeum site were evaluated by statistical analysis using water quality, hydraulic and climate data at Mulgeum site (1998~2008). Before the analysis, control chart method and effect period of typhoon were adopted for improving reliability of the data. After data preprocessing step two methods were used in this study. In method 1, chl-a prediction model was developed using preprocessed data. Another model was developed by Method 2 using significant parameters affecting chl-a after data preprocessing step. As a result of correlation analysis, water temperature, pH, DO, BOD, COD, T-N, NO₃-N, PO₄-P, flow rate, flow velocity and water depth were revealed as significant multiple parameters affecting chl-a concentration. Chl-a prediction model from Method 1 and 2 showed high R² value with 0.799 and 0.790 respectively. Validation for each prediction model was conducted with the data from 2009 to 2010. Training period and validation period of Method 1 showed 20.912 and 24.423 respectively. And Method 2 showed 21.422 and 26.277 in each period. Especially BOD, DO and PO₄-P played important role in both model. So it is considered that analysis of algae occurrence at Mulgeum site need to focus on BOD, DO and PO₄-P.

Key Words : Nakdong River, Chl-a, Multiple parameters, Data Preprocessing, Multivariate Statistics

요약 : 1998년부터 2008년까지 낙동강 물금지점의 수질, 수리 및 기상 인자를 통해 조류발생의 주요 영향인자를 평가하고 chlorophyll-a (chl-a) 예측 모델을 도출하였다. 분석 전 데이터의 신뢰도 향상을 위해 이상치를 제거하는 전처리 과정을 수행하였다. 국내 태풍영향기간을 산정하여 태풍영향기간에 포함된 데이터는 제외를 하였고 관리도 기법을 적용하여 이상치를 제거하였다. 이후 분석과정은 Method 1과 2 두 가지 방법으로 나뉘었다. Method 1은 전처리 과정을 거친 17개 항목의 데이터를 직접 이용하여 chl-a 예측모형을 도출 하였고, Method 2는 전처리 후 주요 영향인자를 도출하여 이를 통해 모형을 개발하였다. Method 2의 상관관계 분석결과 물금지점의 chl-a에 대한 주요 영향인자는 수온, pH, DO, BOD, COD, T-N, NO₃-N, PO₄-P, 유량, 유속 및 수심으로 나타났다. Method 1, 2에 의한 chl-a 회귀모형은 각각 R² 값이 0.799 및 0.790의 높은 유의성을 나타내었고 2009~2010년의 데이터로 검증하였다. Method 1에 의한 chl-a 예측모형의 training period와 validation period의 RMSE 값은 각각 20.912와 24.423으로 나타났으며 Method 2를 통한 모델은 각각 21.422 및 26.277을 나타냈다. 예측 모형 도출 결과 Method 1 및 2 모두 BOD, DO 및 PO₄-P 세 가지 인자가 주요한 역할을 한 것을 볼 때 향후 물금지점의 조류 발생은 BOD, DO 및 PO₄-P를 중점적으로 관리할 필요가 있다고 판단된다.

주제어 : 낙동강, 클로로필-a, 다중영향인자, 데이터 전처리, 다변량 통계기법

1. 서론

낙동강은 총길이가 525 km에 달하는 국내 최대의 하천 중의 하나로써, 4대강(한강, 낙동강, 영산강, 금강) 중에서 연중 부영양화 발생 빈도가 높은 편이다.¹⁾ 1987년 10월 완공된 낙동강 하구언에 의한 낙동강 수계의 체류시간 증가와 기후변화에 의한 기온상승으로 식물성플랑크톤 대량 증식 현상이 저수온기에 자주 관찰되었다.²⁾ 특히 여름철에는 기온상승에 의한 *Microcystis Sp.* 등의 남조류 증식과 녹조 현상으로 인한 시각적인 불편감 및 상수원수에 미치는 악영향 등이 최근까지도 문제가 되고 있다.³⁻⁵⁾ 식물성플랑크

톤의 대량 증식은 정수처리장의 기능 악화와 수중의 용존 산소 감소로 수중 생태계 파괴 및 악취 등을 야기할 수 있는데, 낙동강 하류 지역은 하천수에 의한 용수공급이 90% 이상을 차지하므로^{6,7)} 더 큰 문제가 될 수 있다. 이에 따라 낙동강 하류지역의 조류발생에 관한 연구가 많이 수행되었지만,⁸⁻¹⁰⁾ 대부분의 연구들은 수질인자와 조류발생 사이의 상관관계를 연평균과 월평균 등의 시계열 자료를 통해 해석하는데 집중을 하였다.^{9,11)} 연평균과 월평균 등의 시계열 자료를 통한 분석은 장기적인 수질 및 chlorophyll-a (chl-a) 의 변동 경향 및 패턴에 대한 정보는 제공하였지만, chl-a 와 인자들 간의 구체적인 상관관계를 제시하기에는 한계가

[†] Corresponding author E-mail: big815@pusan.ac.kr Tel: 051-510-7381 Fax: 051-581-5035

있었다. 따라서 일 자료(daily data)를 이용한 chl-a와 복합 영향인자들의 상관관계 분석은 통계적 해석결과 및 조류 발생 예측 모형의 신뢰성을 높일 수 있을 것으로 기대된다. 또한 일 자료를 활용함으로써 우리나라의 여름철에 집중되는 강우 및 태풍 등의 기상현상에 의한 조류발생 예측모형의 왜곡 현상도 저감할 수 있을 것이다. 수질인자 외에도 강우량 및 광량 등과 같은 인자가 chl-a 농도에 미치는 영향을 조사한 연구는 있었지만,^{8,12,13)} 일 자료를 기준으로 수질인자들과 함께 복합적으로 해석한 사례는 거의 없었다. 조류발생은 수질, 수리, 기상 등 다양한 영향인자들에 의해 영향을 받을 뿐만 아니라 이들 영향인자들도 상호 영향을 미치므로 chl-a와의 상관관계 해석 및 조류발생 예측모형 개발을 위해서는 수질, 수리, 기상인자들을 복합영향인자로서 동시에 적용해야할 필요가 있다. 또한 조류발생 제한인자로 잘 알려진 PO₄-P와, 광 에너지 조건과 상관없이 높은 chl-a 농도 관찰이 보고되기도 하였는데, Balbi¹⁴⁾는 다양한 일조 시간에서 chl-a 농도가 관찰되었다고 하였으며, Cao¹⁵⁾은 다양한 빛 강도 조건에서 높은 chl-a 농도가 관찰되었다고 하였다. 또한 Neal¹⁶⁾은 총인 저감시설 설치로 인해 강의 T-P가 현저히 낮아졌지만 이것이 chl-a 농도변화에는 큰 영향을 미치지 않았다고 보고하였다. 따라서 개별인자와 조류발생의 상관관계 해석보다는 복합영향인자들과 조류발생 사이의 상관관계 해석이 더욱 필요할 것이다.

본 연구는 모든 통계적 분석을 위한 데이터를 일자별 자료를 이용하였으며 이를 통해서 낙동강에서의 조류발생에 대한 영향인자를 다변량 통계분석을 이용하여 직접적으로 평가하고자 하였다. 조류 발생 영향인자에 대한 연구를 위한 많은 분석 방법 중 다변량 통계분석은 복합적인 조류발생과 수계의 특성사이의 관계를 평가하고 해석하기 위한 유용한 방법이라고 할 수 있다.¹⁷⁾ 본 연구에서는 pearson 상관계수를 통한 상관관계 분석과 다중회귀분석을 이용하여 물금지역의 조류발생과 영향인자에 대한 연구를 수행하였다. 또한 조류 발생 영향인자로서 수질인자 뿐만 아니라 하천의 수리학적 특성 및 기후조건을 함께 적용한 다중영향인자들과 chl-a 농도 사이의 관계를 평가 하였다.

2. 재료 및 방법

2.1. 대상 지점 및 자료

본 연구의 대상 지점인 물금은 Fig. 1과 같이 낙동강 하류 부산시 취수원 수계(60만 톤/일)에 해당하는 곳이며 밀양강 합류점 이후의 구간으로 낙동강 하구언으로부터 상류 27 km 지점에 위치하고 있다. 2000년대에 들어선 이후 물금지역의 식물성플랑크톤 생체량이 점진적인 감소추세인 반면 우점종의 개체수는 증가하는 추세를 나타내고 있으며 연도별 조류 종의 비율을 볼 때 남조류가 매년 증가하는 추세를 나타내고 있는 지점이다.²⁾ 이를 볼 때 물금 지점은 취수원 수계로서 분석과 관리가 필요한 지점으로써 현재

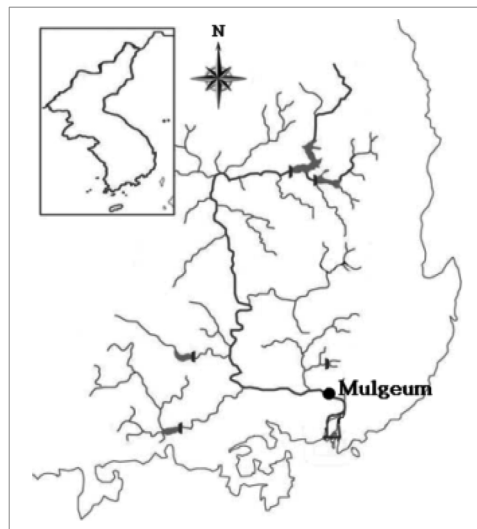


Fig. 1. Location of research area in Nakdong River basin.²¹⁾

다수의 연구기관 및 연구자에 의해 수질 및 조류발생에 대한 모니터링과 연구가 이루어지고 있다.^{2,8-10)}

상관관계 해석 및 조류발생 예측모형 개발에 이용된 자료는 조류의 복합적인 영향인자들을 평가하기 위해 총 17개 항목의 수질(chl-a 포함 11개 항목), 수리(3개 항목) 및 기상자료(3개 항목)를 사용하였으며 그 내용은 Table 1과 같다. 17개 항목의 모든 자료는 물금지역의 동일 날짜 측정자료를 사용하였으며, 1998~2003년 기간은 월 1회 측정자료이고, 2004~2008년 기간은 주 1회 측정자료를 사용하였다. 수질자료는 국립환경과학원 일반측정망 측정자료를 이용하였고, 기상자료는 물금지역과 인접한 부산지역의 측정자료를 이용하였다. 그리고 수리자료는 SWAT분석과 Hec-ras 모형을 이용해 도출된 값을 이용하였다. 수리인자의 경우 측정망 데이터를 이용할 수도 있지만 본 연구는 분석은 같은 일자에 측정된 데이터의 구축을 기본으로 하고 있기 때문에 주 및 월 간격으로 측정된 수질 데이터의 날짜에 맞추어 수리 데이터를 구축하였다. 이와 같은 이유로 측정망의 데이터를 이용하는 것은 한계점이 있었으며 모형을 통해 도출된 필요한 날짜의 수리 데이터를 이용하였다.

1998~2008년 측정 자료를 활용하여 chl-a 농도변화 특성을 분석하였으며, 수질인자 중 조류발생 제한인자로 알려진 PO₄-P는 2001년 3월부터 측정 자료가 존재하므로 2001년 3월부터 2008년 12월까지 17개 항목의 자료를 활용하여 chl-a

Table 1. Data for research (1998~2010)

Classification	Parameters	Resource
Water quality	Water temperature, pH, DO, BOD,	National Institute of Environmental Research
	COD, T-N, NO ₃ -N, NH ₄ -N, T-P, PO ₄ -P('01/03-'08/12), chl-a	
Hydraulic	Flow rate, flow velocity, water depth	SWAT analysis, Simulation result of Hec-ras model
Climate	Precipitation, insolation, wind velocity	The meteorological Administration

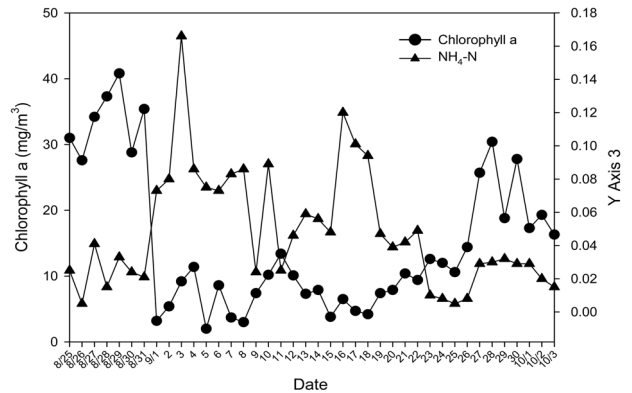
와 복합영향인자들의 상관관계 분석 및 chl-a 농도 예측 모형을 개발하였다. 그리고 예측모형의 검증은 2009~2010년의 측정 자료를 이용하여 실시하였다.

2.2. 통계적 분석 방법

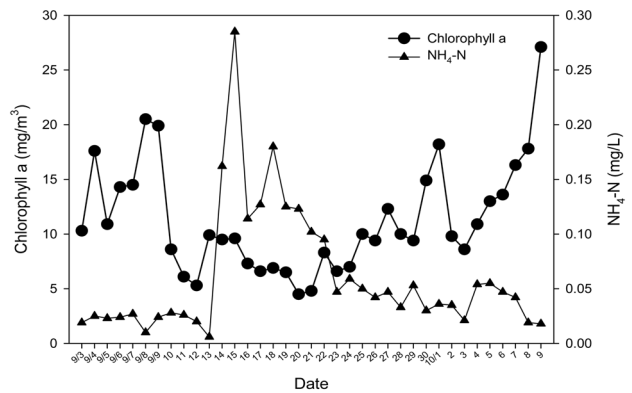
본 연구의 chl-a와 복합 영향인자들의 상관관계 분석 및 chl-a 예측모형 개발은 SPSS ver.21을 이용하여 Fig. 2에 제시된 바와 같이 3단계로 구분되어 수행되었다. Step 1에서는 자료의 신뢰성 향상을 위한 전처리(data preprocessing)를 실시하였고, Step 2에서는 Method 1과 2를 이용한 상관관계 분석 및 예측모형 개발을 실시하였다. 그리고 Step3에서는 예측모형의 검증을 실시하였다.

2.2.1. Step 1 : 데이터 전처리

본 연구에서는 통계적 분석에 앞서 자료의 전처리를 실시하였다. 관리도 기법(Congrol Chart Method)을 적용하여 각 변수들의 평균값에서 $\pm 3\sigma$ 의 범위내의 자료들을 대상으로 분석을 실시하였으며,¹⁸⁾ 또한 태풍에 의한 수질과 수리 자료의 왜곡현상을 배제하기 위해 태풍영향기간을 산정하고 이 기간의 측정자료를 제외하였다. 낙동강 수계에 대한 태풍영향기간은 ‘국가태풍센터’의 태풍 상륙 및 소멸시기 정보를 이용하여 태풍 상륙 이후 물금지역 수질 농도(부산시 상수도사업본부 수질연구소 매일 측정자료) 중 chl-a와 NH₃-N의 농도 변화 및 태풍상륙 이전 수준의 농도 회복을 기준으로 산정하였다. 국내 상륙 태풍 중에서 가장 큰 규모로 영향을 미친 2002년의 태풍 ‘RUSA’와 2003년의 태풍



(a) Conc. of chl-a and NH₄-N in period of typhoon ‘RUSA’



(b) Conc. of chl-a and NH₄-N in period of typhoon ‘MAEMI’
Fig. 3. Conc. of chl-a and NH₄-N in period of typhoon ‘RUSA’ and ‘MAEMI’ at Mulgeum.

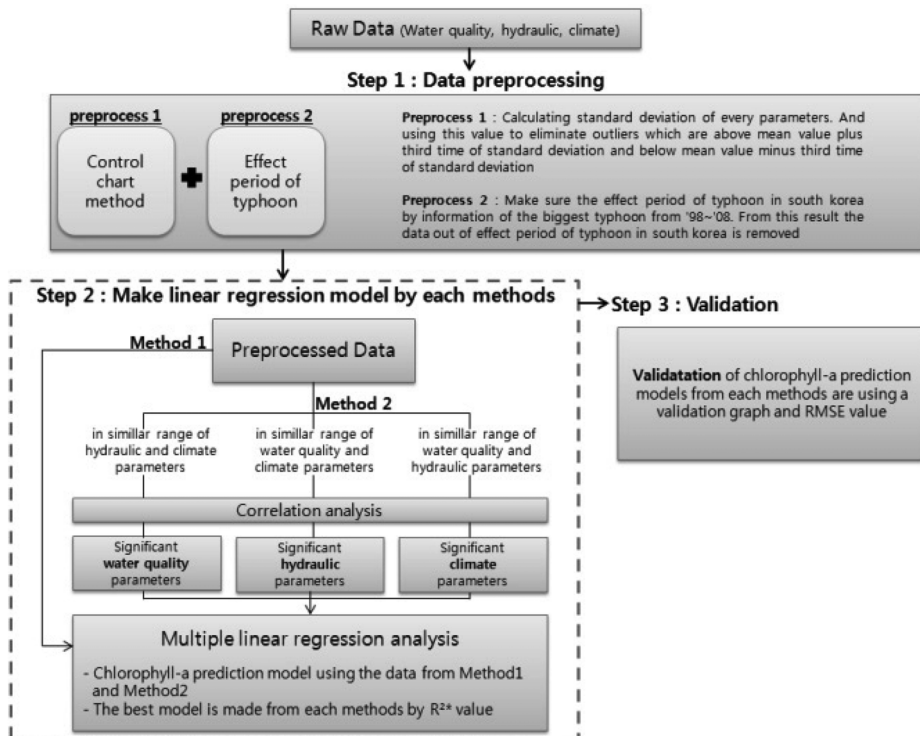


Fig. 2. The proposed flow chart for determining statistically optimized regression variables and developing the chlorophyll-a prediction model by multiple linear regression.

‘MAEMI’를 대상으로 태풍영향기간을 산정하였다. 비교결과(Fig. 3)를 보면 2002년 태풍 ‘RUSA’의 경우, 9월 1일을 기점으로 chl-a 농도가 35.4 mg/m³에서 3.2 mg/m³으로 크게 감소하였는데, 이것은 태풍이 동반한 폭우로 인한 물금 지역 유입유량 증대로 조류의 플러싱효과가 일어났기 때문으로 판단되며 태풍발생 이전의 농도 수준(35.4 mg/m³)으로 회복되기까지는 약 25일이 소요되었다. NH₄-N 역시 9월 1일을 기점으로 비점오염물질 유입으로 0.021 mg/L에서 0.073 mg/L로 농도가 대폭 증가하였으며 태풍상륙 이전 수준의 농도를 나타내는데 9월 22일까지 약 22일 정도의 기간이 소요되는 것으로 나타났다.

태풍 ‘MAEMI’의 경우는 2003년 9월 6일 우리나라에 상륙하여 소멸되기까지 4일이 경과되었고 9월 10일 이후 chl-a의 농도가 19.9 mg/m³에서 8.6 mg/m³으로 감소하였으며 이후 태풍이전 농도로 회복되는데 22일이 소요되었다. NH₄-N 농도도 9월 13일 이후 0.006 mg/L에서 0.162 mg/L로 크게 증가한 이후 농도가 이전 수준으로 회복되는데 19일 정도가 걸리는 것으로 나타났다.

각 태풍 기간 전, 후 chl-a의 평균농도를 보면 ‘RUSA’ 기간은 태풍 전 일주일간의 chl-a는 33.6 mg/m³이었고 ‘MAEMI’ 때에는 14.9 mg/m³이었다. NH₄-N의 경우는 각각 0.021, 0.022 mg/m³이었고 태풍 영향기간 중 chl-a 및 NH₄-N의 평균농도는 각각 17.3 mg/m³ 및 0.076 mg/L로 나타났다. 그리고 영향기간 22일 내지 25일이 지난 후 최초 일주일간의 평균농도는 chl-a가 각각 30.41, 18.3 mg/m³, NH₄-N이 0.023, 0.006 mg/L로 영향기간 이전 수준의 농도로 회복이 되었음을 알 수 있었다. 이러한 결과는 선행연구자들인 Wenrui 등(2011)이 미국의 Pensacola 만(Bay)을 대상으로 hurricane 이 chl-a에 미치는 영향을 분석한 연구결과와도 대략적으로 일치하였다.

2.2.2. Step 2 : 전처리된 데이터를 이용한 Chlorophyll-a 예측 모델

Step 2에서는 chl-a 예측 모델 개발을 위해 2가지 방법이 사용되었다. Method 1은 전처리 과정을 거친 데이터들을 다중회귀분석에 적용시켜²⁰⁾ chl-a를 예측하는 모델을 도출하는 것이다. 이와 달리 Method 2는 전처리 과정을 거친 데이터를 이용해서 수질, 수리, 기상인자 각 종류별로 chl-a 농도에 주요한 영향을 끼치는 인자를 도출한 후 도출된 주요 영향인자를 다중회귀분석에 적용시켜 모델을 도출 하였다. 주요 영향인자 도출 과정에서 수리중요인자 도출을 위해 수리, 기상인자의 데이터 범위에 제한을 두어 수리, 기상인자의 영향을 배제하고 수질인자와 chl-a 농도의 상관성을 평가하였다. 마찬가지로 수리 주요 영향인자를 구할 때는 수질, 기상인자의 범위를 제한, 기상 주요 영향인자의 경우는 수질, 수리인자의 범위를 제한하고 분석을 진행하였다.

2.2.3. Step 3 : 검증

도출된 두 예측 모델을 대상으로 결정계수(R²)값을 통해

모델의 유의성을 평가하고 모델 도출에 사용된 이후의 데이터(2009~2010)를 통해 모델의 검증을 실시하였다. 또한 RMSE 값을 통해 두 모델의 정확도²⁰⁾를 평가하였다.

3. 결과 및 고찰

3.1 Chlorophyll-a의 변동특성

물금 지점에서 1998년부터 2008년까지 11년간의 월 평균 chl-a 농도를 평가한 결과 계절별로 농도 범위가 뚜렷하게 나뉘었다(Fig. 4). 각 계절별 chl-a 평균 농도는 봄철이 59.51 mg/m³, 여름철이 26.26 mg/m³, 가을철이 33.27 mg/m³, 겨울철이 103.39 mg/m³로 겨울철이 다른 계절에 비해서 높은 chl-a 평균 농도를 보였다.⁷⁾ 조류 예·경보제의 chl-a 기준 농도와 비교했을 때 봄, 여름, 가을은 조류경보의 등급에 해당하며 겨울철은 조류 대발생의 등급에 해당하는 chl-a 평균 농도를 보였다. 여름철은 chl-a의 농도는 낮지만 녹조 현상과 남조류에 의한 독소가 주로 문제가 되며 현재 여름의 녹조류 개체수는 증가하는 추세를 보이고 있다.²⁾ 겨울철은 높은 chl-a 농도 때문에 정수처리에 있어 악영향을 미친다. 겨울철은 낮은 수온으로 인해 조류 등 미생물이 성장하기에는 좋지 않은 조건이지만, 평균 chl-a의 농도가 높은 이유는 낮은 온도에서 최적의 성장 범위를 보이는 저온성 규조류인 *Stephanodiscus sp.*의 대량 증식 현상에 의한 것으로 보고되었다.^{8,21,22)} 반면 여름철은 조류의 성장에 겨울철에 비해 더욱 적절한 수온임에도 불구하고 낮은 chl-a 평균농도를 보였다. 따라서, 봄, 여름, 가을 및 겨울의 chl-a 농도가 다르게 나타나는 것은 각 계절별로 나타나는 우점종의 차이뿐만 아니라 수질, 수리 및 기상 인자 등 chl-a에 영향을 미치는 인자들의 복합적 특성에 기인하는 것으로 판단된다.

3.2. Data 전처리 (Step 1)

Data의 이상치를 제거하기 위한 과정으로 관리도 기법과 국내 상륙태풍에 의한 영향기간을 적용시킨 데이터 전처리

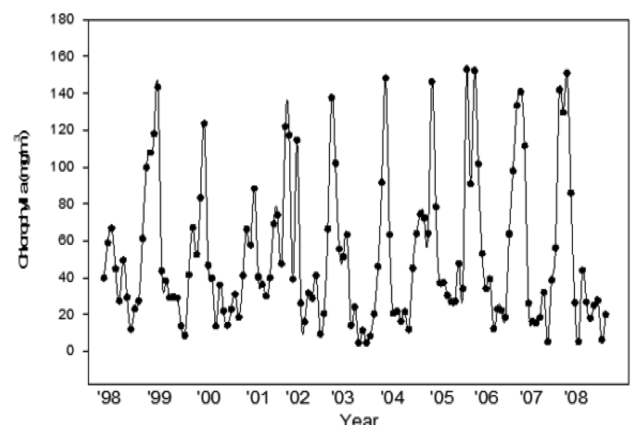


Fig. 4. Monthly variation of chlorophyll-a concentration from 1998 to 2008 at Mulgeum.

Table 2. Correlation analysis between chl-a and non-preprocessed data

	W.T.	pH	DO	BOD	COD	T-N	NO ₃ -N	NH ₄ -N
R	-0.698	0.680	0.798	0.826	0.656	0.572	0.502	-0.053
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.380

	T-P	PO ₄ -P	Flow rate	Flow velocity	Water depth	Wind velocity	Precipitation	Insolation
R	0.114	-0.654	-0.111	-0.164	-0.216	0.013	-0.096	-0.208
p	0.056	0.000	0.064	0.006	0.000	0.835	0.111	0.000

Table 3. Correlation analysis between chl-a and preprocessed data

	W.T.	pH	DO	BOD	COD	T-N	NO ₃ -N	NH ₄ -N
R	-0.696	0.675	0.793	0.816	0.731	0.586	0.513	-0.106
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.092

	T-P	PO ₄ -P	Flow rate	Flow velocity	Water depth	Wind velocity	Precipitation	Insolation
R	0.206	-0.650	-0.176	-0.172	-0.159	-0.026	-0.093	-0.204
p	0.001	0.000	0.005	0.006	0.012	0.681	0.143	0.001

(Step 1)를 수행하였다. Step 1의 수행 전 Raw data를 이용해 각 인자들과 chl-a와의 상관관계를 pearson 상관계수를 통해 분석한 결과 수온, pH, DO, BOD, COD, T-N, NO₃-N, PO₄-P, Flow rate, Water depth 및 Insolation이 chl-a 농도와 유의한 상관관계 (p<0.05)를 나타냈다(Table 2). 그중 BOD, DO, 수온, pH, COD 및 PO₄-P가 상대적으로 높은 상관관계를 나타냈다. 반면 Step 1을 거친 data를 대상으로 chl-a와 각 인자 간의 상관관계를 pearson 상관계수를 통해 평가한 결과는 전처리 전과 달리 T-P와 Flow rate 인자가 추가적으로 chl-a와 유의한 상관관계를 나타냈다(Table 3). 다른 유의한 상관관계를 나타냈던 인자들의 상관계수 값은 큰 차이를 보이지 않았으며 COD가 0.656에서 0.731로 증가, Water depth가 -0.216에서 -0.159로 감소한 결과를 보였다. 결과적으로 집중호우, 태풍 및 측정오류 등이 다른 인자들 보다 T-P와 Flow rate의 데이터에 더 크게 영향을 미쳤다고 평가된다. 이후 단계의 분석은 전처리 과정을 거친 data를 통해 실시되었다.

3.3. Method 1을 이용한 chl-a 예측 모형(Step 2)

전처리 과정을 거친 수질, 수리 및 기상인자 16개 항목에 대한 데이터를 이용하여 chl-a 농도를 종속변수로 하는 다중선형회귀분석을 실시하였다. 분석과정에서 유의한 변수들을 선별하기 위해 종속변수에 영향을 줄 것으로 생각되는 변수들을 하나씩 추가하며 매 단계별로 변수가 추가될 때 중요도가 상실된 변수는 다시 제거하게 되는 Stepwise 변수 선택법을 사용하였다. 도출된 모델에서 독립변수들 간의 상관성을 가지는 다중공선성이 유량, 유속, 수심의 수리인자에서 나타났는데(VIF value > 10)²⁰⁾ 이것은 수심, 유속 인자가 유량의 결과로부터 모의된 값이므로 이러한 결과가 나타난 것으로 판단된다. 결과적으로 수리인자 중 유량만을

Table 4 Result of multiple linear regression by Method 1

No.	Selected parameters in regression model	R ^{2*}
1	BOD	0.665
2	BOD, DO	0.754
3	BOD, DO, PO ₄ -P	0.781
4	BOD, DO, PO ₄ -P, T-P	0.797
5	BOD, DO, PO ₄ -P, T-P, Insolation	0.799

Table 5. Result of the selected multiple linear regression model of chl-a by Method 1

Parameters	Coefficient	Beta coefficient	p-value	VIF
Constant	-36,277			
BOD	17,742	0.378	0.000	2,516
DO	4,775	0.309	0.000	2,408
PO ₄ -P	-572,959	-0.314	0.000	2,317
T-P	259,966	0.158	0.000	1,580
Insolation	-0.405	-0.061	0.037	1,052

포함시킨 chl-a 예측 모델은 총 5개가 도출되었다(Table 4). 이 중 BOD, DO 및 PO₄-P 3가지 인자만을 통해 결정계수가 0.781을 나타내는 것을 볼 때, BOD, DO 및 PO₄-P가 chl-a 농도 예측에 가장 큰 역할을 하는 것으로 판단된다.

도출된 5개의 모델 중 R^{2*}값이 가장 높게 나타난 5번째 모델은 아래의 식과 같으며 총 5개의 변수를 포함하고 있다.

$$Chl.a = -36.277 + 17.742BOD + 4.775DO - 572.959PO_4P + 259.966TP - 0.405Insolation \quad (R^{2*} = 0.799) \quad (1)$$

모델의 p-value와 VIF값을 통해 선택된 변수들이 모두 유의하며 다중공선성을 가지지 않는 것을 확인할 수 있다(Table 5). 또한 모델 내의 종속변수에 대한 상대적인 영향력의 크기를 나타내는 Beta계수는 BOD와 PO₄-P 그리고 DO순으로 높게 나타났다. 따라서 만들어진 모델의 종속변수인 chl-a 농도에 크게 영향을 미치는 변수로 BOD와 PO₄-P 그리고 DO를 말할 수 있으며 T-P와 Insolation의 경우는 chl-a의 농도에 영향을 미치는 유의한 변수이나 다른 인자들에 비해 낮은 Beta계수 값을 나타내고 있다.

3.4. Method 2를 이용한 chl-a 예측 모형(Step 2)

3.4.1. 수질 영향인자

수질, 수리 및 기상인자 중 수질인자를 대상으로 chl-a 농도 변화에 대한 주요 인자를 평가하고 나머지 인자에 대한 영향을 최소화 하여 분석을 실시하였다. 분석 과정에서 수리인자의 경우 유량이 조류 농도에 큰 영향을 미치지 않게 하기 위해 1998년부터 2008년 사이의 평균 저수량(127.01 m³/s)을 산정하여 유량이 평균 저수량 이하인 날을 선정하였다. 이때, 유속과 수심은 유량에 의해 모의된 데이터로 유량의 구분을 통해 함께 고려되었다. 그리고 기상인자 중 일사량은 11년간의 평균치인 14.4 MJ/m²보다 ±σ (7.1 MJ/m²)의 양에

Table 6. Seasonal value of pearson coefficient of correlation between water quality data and chl-a

		W.T.	pH	DO	BOD	COD	T-N	NO ₃ -N	NH ₄ -N	T-P	PO ₄ -P
R	Spring, Autumn	-0,651	0,637	0,777	0,847	0,823	0,430	0,480	-0,143	0,057	-0,852
	Summer	0,102	-0,199	-0,261	0,749	0,566	-0,641	-0,573	-0,392	-0,395	-0,607
	Winter	-0,239	0,744	0,213	0,786	0,450	0,765	0,493	0,002	0,195	-0,694

해당하는 구간으로서 광량이 너무 많지도 적지도 않게 하여 조류 발생에 큰 영향을 주지 않도록 제한하였다. 또한 강수량은 일일 강수량이 0 mm로 강수에 대한 직접적인 영향이 없는 날을 선택하였다. 이러한 과정들을 통해 수리 및 기상 인자의 영향을 배제하고 수질 인자들과 chl-a와의 상관관계를 분석하였다. 이 때 분석 중 계절별로 온도의 조류 발생에 대한 영향 인자들을 평가하기 위해 겨울철, 여름철 및 나머지 계절을 구분하여 분석을 실시하였다. 상관관계 분석 결과 영향인자는 p-value가 0.05 이하를 나타내는 인자를 대상으로 하였다. 분석 결과 봄, 가을철은 수온, pH, DO, BOD, COD, T-N, NO₃-N과 PO₄-P, 여름철은 BOD 그리고 겨울철은 pH, BOD 및 T-N이 chl-a와 유의한 관계를 보였다(Table 6). BOD는 측정값에 chl-a가 포함될 수 있으므로 이를 주요 영향인자라고 판단하기는 어려우며 이에 대한 연구가 더 필요한 것으로 판단된다. 또한 계절별로 chl-a의 주요 영향인자가 다르게 나타났는데, 이것은 계절에 따라 서로 다른 최적 성장범위를 가지는 우점종이 존재하기 때문으로 보고되어진다.²³⁾ 따라서 수리, 기상 인자의 범위를 제한했을 때 chl-a 농도 변화에 영향을 미치는 인자는 수온, pH, DO, BOD, COD, T-N, NO₃-N 및 PO₄-P로 나타났으며 이중 pH, BOD 과 T-N은 3계절 이상을 chl-a와 유의한 상관성을 나타내었다.

3.4.2. 수리 영향인자

수질인자 중 식물성플랑크톤이 이용할 수 있는 용존성 무기염류 형태의²³⁾ PO₄-P, NO₃-N, NH₄-N의 농도 제한 범위는 물금지점에 존재하는 자료수집기간 동안의 T-N, T-P, PO₄-P, NO₃-N, NH₄-N의 평균 농도와 조류의 생장에 필요한 최소한의 T-N, T-P 농도(0.2, 0.02 mg/L)²⁴⁾ 사이의 비(ratio)를 이용해 추정하였다. 그 결과 조류 생장에 대한 최소농도는 PO₄-P, NO₃-N, NH₄-N이 각각 0.009, 0.133, 0.007 mg/L로 나타났다. 이를 물금지점 전체 데이터와 비교한 결과 전체 기간에 걸쳐 NO₃-N에 대한 최소조류증식 농도를 넘어서고 있는 상태이며, 이를 통한 조류발생 가능성은 항상 잠재적으로 지니고 있는 상태로 판단된다. PO₄-P와 NH₄-N은 산정된 최소조류증식에 대한 농도보다 낮은 날의 데이터는 분석 데이터에서 제외하였다. pH 5 미만의 산성화된 수중에서 다른 영양소가 풍부하면 CO₂같은 무기탄소가 생장을 제한하게 되는데²⁵⁾ 수중의 CO₂에 대한 데이터는 분석에 이용하지 않기 때문에 pH가 5 이상인 범위를 선정하였으며, 광합성 효소의 정상적인 활동을 위해 중성범위를 크게 벗어나지 않도록 pH가 9 이상인 데이터는 배제하고 분석을 진행하였다. DO는 국내 하천생활환경기준의 보통 수준 이상에 해당하는 기간을 분석에 이용하였으며 BOD와 COD

Table 7. Seasonal value of pearson coefficient of correlation between hydraulic condition and chl-a

		Flow rate	Flow velocity	Water depth
R	Spring, Autumn	-0,267	-0,277	-0,269
	Summer	-0,084	-0,095	-0,099
	Winter	0,479	0,492	0,510

는 chl-a의 농도를 포함하므로 제한 범위를 두지 않았다. 기상인자는 수질영향인자 도출의 경우와 같은 범위인 일 강수량 0 mm 및 평균 일사량 14.4 MJ/m²의 ±σ(7.1 MJ/m²)범위 안의 자료를 선택하였다. 분석결과는 Table 7과 같으며 봄, 가을, 겨울철의 경우 수리인자 모두 chl-a와 유의한 상관관계를 보였다(p<0.05). 여름철의 경우 집중호우나 댐 방류 등에 의한 플러싱효과 등 태풍의 영향만이 아니라 일시적인 기상으로 인한 큰 영향으로 잦은 변화를 보이며 특히 주 1회 측정 데이터를 통해서도 기상현상들의 해석에 한계점이 있는 것으로 판단이 된다. 따라서, 향후 여름철의 기상현상과 댐 방류 및 일(daily)데이터의 보완을 통해 보다 더 신뢰도를 향상시킨 분석이 필요할 것으로 사료된다.

3.4.3. 기상 영향인자

앞서 분석에 이용한 수질, 수리 인자들의 범위를 제한하고 chl-a 농도와 기상인자들 간의 상관도를 분석한 결과 기상인자들은 모두 p-value가 0.05 이상으로 chl-a와 유의한 상관관계를 보이지 않았다(Table 8). 하지만 일사량은 조류의 성장을 위한 광합성에 기본적으로 필요한 인자로서²³⁾ 향후 새로운 분류기준을 통해 평가해볼 필요성이 있다. 강수량 역시 유량에 직접적으로 영향을 미치는 인자이므로 본 연구에 사용된 한 지점이 아닌 연구 대상지점 모든 상류지역 강수량을 적용한 유역 평균 강수량을 산정하여 이를 적용할 필요가 있는 것으로 판단된다.

3.4.4. 주요영향인자를 통한 chl-a 예측 모델

Chl-a 농도변화에 대한 수질 영향인자는 수온, pH, DO, BOD, COD, T-N, NO₃-N, PO₄-P, 유량, 유속 및 수심으로 나타났다. 따라서 낙동강 하류 물금지점은 기상 보다 수질,

Table 8. Seasonal value of pearson coefficient of correlation between climate condition and chl-a

		Precipitation	Insolation
R	Spring, Autumn	0,107	0,033
	Summer	0,305	0,370
	Winter	0,057	0,390

Table 9. Result models of multiple linear regression by Method 2

No.	Selected parameters in regression model	R ^{2*}
1	BOD	0.665
2	BOD, DO	0.754
3	BOD, DO, PO ₄ -P	0.781
4	BOD, DO, PO ₄ -P, COD	0.787
5	BOD, DO, PO ₄ -P, COD, Water temperature	0.790

수리인자가 chl-a 영향인자로 더 유력하며 상관계수 값이 훨씬 높게 나타나는 수질 인자가 더 유의한 인자로 추측된다. Chl-a 농도변화가 기상인자와는 상관관계를 보이지 않았지만 향후 인자별로 데이터의 보완을 통한 증·상류 지점과의 비교연구가 필요할 것으로 판단된다. 도출된 주요영향인자인 수온, pH, DO, BOD, COD, T-N, NO₃-N, PO₄-P, 유량은 chl-a의 예측 모형의 독립변수로 사용되었다. 유속 및 수심은 유량에 의해 도출된 값이므로 다중공선성을 없애기 위해 변수로 입력하지 않았다. 입력된 변수들을 이용하여 만들어진 모형의 결과는 Table 9에 나타내었다. T-N의 데이터는 회귀모형 도출 과정의 stepwise 변수선택법 적용과정에서 유의하지 않은 변수로 판단되어 변수로 선택되지 않았다(p>0.05).

도출된 모형중 가장 높은 결정계수를 가지는 회귀모형은 아래와 같으며 R^{2*} 값이 0.790으로 높은 유의성을 나타내었다.

$$\text{Chl.a} = -28.523 + 13.095\text{BOD} + 4.570\text{DO} - 403.782\text{PO}_4\text{P} + 5.929\text{COD} - 0.576\text{Water Temp.} \quad (R^{2*} = 0.790) \quad (2)$$

이 때, stepwise 변수선택법에 의해 도출된 모형을 살펴보면 Method 1과 마찬가지로 BOD, DO, PO₄-P 만으로 모형의 결정계수가 0.781까지 결정되는 것을 알 수 있다. 결국 Method 2를 통한 분석에서도 Method 1 같은 결과를 나타내는 것을 볼 때, 낙동강 하류 물금지점은 BOD, DO, PO₄-P가 chl-a 농도 예측값을 생성하는데 가장 큰 역할을 하고 있음을 알 수 있다. Table 10은 Method 2를 통한 모형의 모든 변수들이 유의하며 다중공선성을 보이지 않음을 보여준다. 그리고 Beta계수를 통해 BOD, 수온, PO₄-P의 순서로 도출된 모형의 종속변수인 chl-a 농도 예측에 큰 영향을 미치며 pH의 경우 모형에 유의한 독립변수로 판단되어 모형에 변수로 포함이 되었지만 종속변수의 변화에 큰 영향을 주지는 않는 것으로 판단된다.

Table 10. Result of the selected multiple linear regression model of chl-a by Method 2

Parameters	Coefficient	Beta coefficient	p-value	VIF
Constant	-28.523			
BOD	13.095	0.279	0.000	4.326
DO	4.570	0.296	0.000	3.643
PO ₄ -P	-403.782	-0.222	0.000	1.599
COD	5.929	0.158	0.002	3.032
Water temp.	-0.576	-0.103	0.043	3.046

3.5. Chl-a 예측 모형의 검증(Step 3)

Method 1과 2에 의해 도출한 chl-a 예측 모형의 training 기간(2001~2008년)과 validation 기간(2009~2010년) 동안의 chl-a 농도 예측값과 실제 측정값은 Fig. 5와 6을 통해 나타났다. Training은 데이터를 이용해서 chl-a 예측 모형을 도출한 것을 이야기하며 validation은 training에 의해 도출된 모형을 training 기간이 아닌 다른 기간('09~'10)의 데이터를 사용하여 도출된 모형이 잘 들어맞는지 확인하는 과정을 의미 한다. Fig. 5~6을 통해 validation기간의 실제 측정값과 예측값의 차이 정도를 비교할 수 있다. 그리고 RMSE값을 이용하여 두 모형의 정확도를 평가한 결과는 Table 11을 통해 알 수 있다. Method 1에 의해 도출된 모형의 training 기간의 RMSE값이 20.912, validation 기간의 RMSE값이 24.423로 Method 2에 의해 도출된 모형(training 기간의 RMSE 값이 21.422, validation 기간의 RMSE값이 26.227)과 거의 유사한 값을 나타냈다. RMSE 값을 이용한 정확도를 볼 때 Method 1을 이용한 chl-a 예측 모형이 조금 더 높은 정확도를 나타낸 것을 알 수 있다. Method 2를 통해 도출된 모형은 Method 1을 통해 도출된 모형과 같은 독립 변수의 개수를 통해 유사한 수준의 모형이 도출되었다. 하지만 Method 2를 통한 모형 내부의 독립변수의 변화가 존재하는 것을 알 수 있었다. 따라서, chl-a의 예측모형 도출에 있어서 주

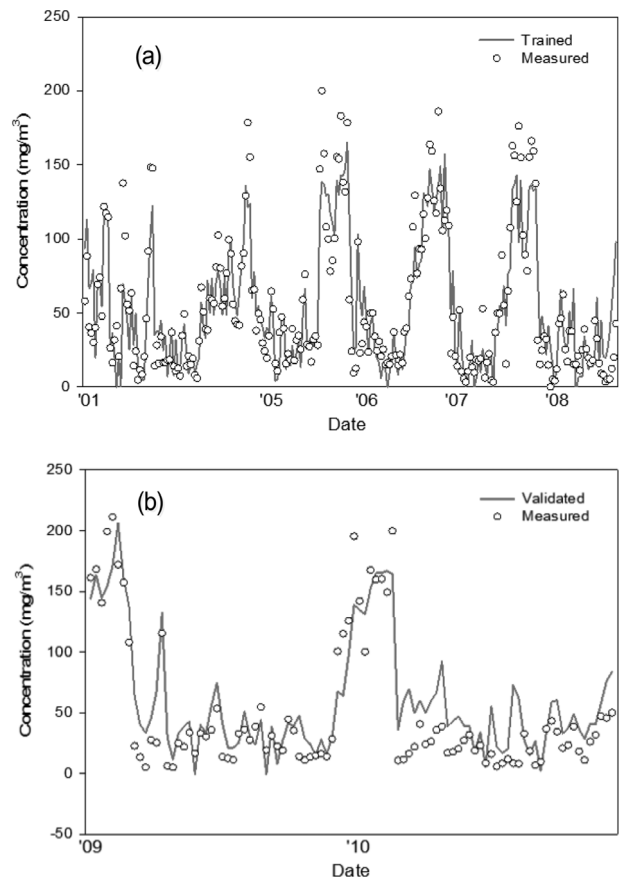


Fig. 5. Training and validation results of MLR by Method 1 for chl-a prediction [(a):training result, (b): validation result].

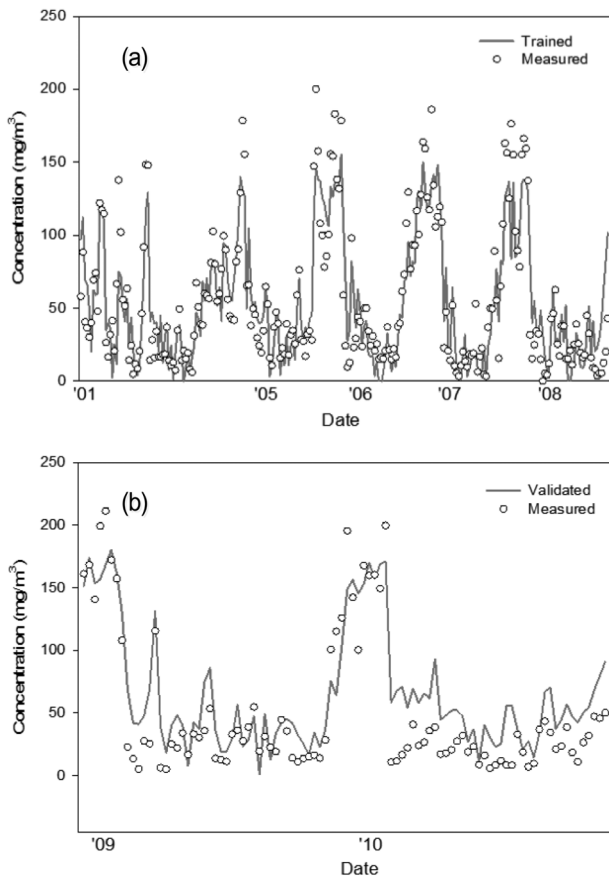


Fig. 6. Training and validation results of MLR from selected parameters by method 2 for chl-a prediction [(a): training result, (b): validation result].

Table 11. RMSE value of both regression model for prediction chl-a

MLR model		RMSE
Model 1 (by Method 1)	training	20,912
	validation	24,423
Model 2 (by Method 2)	training	21,422
	validation	26,227

요 영향인자를 도출하는 방법(Method 2)을 이용하는 것은 데이터의 지속적인 구축과 더 명확한 인자들의 제한범위가 함께 병행될 필요가 있다고 판단된다.

4. 결론

본 연구에서는 낙동강 하류 물금 지점의 조류발생 주요 영향인자 평가 및 chl-a 농도 예측 모형 개발을 위해 수질, 수리, 기상 인자를 복합적으로 평가하였으며 분석결과는 다음과 같다.

1) 데이터 전처리 과정으로 관리도 기법(Control Chart Method)과 태풍 영향기간 산정을 통해 낙동강 하류 물금 지점

의 데이터들의 이상치를 제거한 결과 T-P와 Flow rate가 chl-a와 유의한 상관관계를 추가적으로 나타냈다.

2) 1998년부터 2008년까지 계절별로 chl-a 농도특성을 평가한 결과, 겨울철의 chl-a 농도가 가장 높은 반면 여름철의 chl-a 농도가 가장 낮은 것으로 나타났다. 이는 겨울에 최적 성장범위를 가지는 저온성 규조류의 우점, 증식 현상과 우리나라 기후 특성상 여름에 집중된 강우로 인한 플러싱 효과 때문인 것으로 판단된다.

3) 수질, 수리 및 기상인자와 chl-a 농도 사이의 상관관계 분석결과 낙동강 하류의 물금지점에서 chl-a에 대한 주요 영향인자는 수질인자인 것으로 추측되며 수질인자 중 수온, pH, DO, BOD, COD, T-N, NO₃-N 및 PO₄-P 8가지가 chl-a에 유의한 상관관계가 있는 것으로 나타났다. 수리인자는 chl-a와 유의한 상관관계를 나타내지만 낮은 상관계수 값을 보였다.

4) Method 1에 의한 chl-a 예측 모형은 0.799의 높은 결정계수(R²)값을 보였으며, training 기간과 validation 기간의 RMSE 값은 각각 20.912 및 24.423으로 나타났다. Method 2에 의해 도출된 모형 역시 R²=0.790의 높은 결정계수 값을 보이며 RMSE 값은 각각 21.422와 26.227을 보였다. 두 방법의 모형 모두 높은 유의성을 나타냈으나 Method 1에 의한 모형이 더 높은 정확도를 보였다.

5) Method 1과 Method 2를 통한 chl-a 예측 모형을 볼 때 물금지점에서 BOD, DO, PO₄-P 인자가 앞서 도출된 주요 영향인자 가운데서도 chl-a 농도를 예측하는데 가장 큰 역할을 한 것을 알 수 있다.

Acknowledgement

이 연구는 2014년 정부(환경부)의 재원으로 부산녹색환경지원센터의 지원을 받아 수행된 연구(14-2-30-33)이며, 한국연구재단 BK21플러스 사업의 일환으로 수행된 연구결과이다(21A20132012304).

KSEE

References

- Hwang, D. J., Lee, H. J., Yoon, J. S., Heo, S. N., Im, T. H., Kwon, Y. H., Shin, C. K. and Kim, H. W., "Study on Ecological Characteristic of Algal Species Forming Bloom during Spring-Winter Season in the Nakdong River(II)," Nakdong River Environment Research Center and Suncheon national University(2007).
- Son, H. J., "Changes of Dominant Phytoplankton community in Downstream of the Nakdong River: From 2002 to 2012," *J. Korean Soc. Environ. Eng.*, **35**(4), 283-293(2013).
- Lee, C. S., Ahn, C. Y., La, H. J., Lee, S. H. and Oh, H. M., "Technical and Strategic Approach for the Control of cyanobacterial Bloom in Fresh Waters," *Korean J. Environ. Biol.*, **31**(4), 232-242(2013).
- Kim, H. B., Park, H. K., Shin, K. D. and Moon, J. S., "The

- Characteristics of Toxin Production in the Korean Toxic Cyanobacteria,” *J. Korean Soc. Water Qual.*, **26**(5), 834~840 (2010).
5. Lee, K. L., Jheong, W. H., Kim, J. M., Kim, Y. S., Choi, H. J. and Kim, H. S., “Seasonal Variations of cyanobacterial Toxins (microcystins) in Yeongchun Reservoir,” *Korean J. Limnol.*, **41**(2), 264~274(2008).
 6. Nakdong River Basin Environmental Office, “Analysis of hydraulic characteristics and development of hydraulic modeling technology for total water pollution management for the Nakdong river,” Final report(2006).
 7. Park, H. K., “Survey Method relating Freshwater Phytoplankton for the Management of Water Resources,” *J. Korean Soc. Environ. Eng. Special Feature*, **29**(6), 593~609(2012).
 8. Joung, S. H., Park, H. K., Lee, H. J. and Lee, S. H., “Effect of Climate Change for Diatom Bloom at Winter and Spring Season in Mulgeum Station of the Nakdong River, South Kroea,” *J. Korean Soc. Water Environ.*, **29**(2), 155~164 (2013).
 9. Son, H. J., “Long-Term Variations of Phytoplankton Biomass and Water Quality in the Downstream of Nakdong River,” *J. Korean Soc. Environ. Eng.*, **35**(4), 263~267(2013).
 10. Chung, J. M., Choi, C. M. and Moon, S. G., “The Structure of Phytoplankton community in the Middle-Lower Part of the Nakdong River,” *J. Korean Environ. Sci. Soc.*, **10**(1), 41~45(2001).
 11. Cho, K. H., Kang, J. H., Ki, S. J., Park, Y. E., Cha, S. M. and Kim, J. H., “Determination of the optimal parameters in regression models for the prediction of Chlorophyll-a : A case study of the Yeongsan Reservoir, Korea,” *Sci. Total Environ.*, **407**, 2536~2545(2009).
 12. Kim, J. A., Kim, G. A., Yun, C. J., Park, H. G., Jung, E. Y., Cha, D. J., Choi, J. T. and Son, H. J., “A study on the Correlation between Odorous Compounds, Actinomycetes and Algae in Drinking Water Source of Nakdong River,” *J. Korean Soc. Environ. Eng.*, **35**(3), 213~219(2013).
 13. Cao, C., Zheng, B., Chen, Z., Huang, M., Zhang, J., “Eutrophication and algal blooms in channel type reservoirs : A novel enclosure experiment by changing light intensity,” *J. Environ. Sci.*, **23**(10), 1660~1670(2011).
 14. David, M. B., “Suspended chlorophyll in the River Nene, a small nutrient-rich in eastern England: long-term and spatial trends,” *Sci. Total Environ.*, **251/252**, 401~421(2000).
 15. Cao, C., Zheng, B., Chen, Z., Huang, M. and Zhang, J., “Eutrophication and algal blooms in channel type reservoirs: A novel enclosure experiment by changing light intensity,” *J. Environ. Sci.*, **23**(10), 1660~1670(2011).
 16. Neal, C., Hilton, J., Andrew, J. W., Neal, M. and Wickham, H., “Chlorophyll-a in the rivers of eastern England,” *Sci. Total Environ.*, **365**, 84~104(2006).
 17. Jeong, Y. H., Kim, H. S. and Yang, J. S., “Statistical Analyses of Long-Term water quality variation in the Geumgang-Reservoir; Focused on the TP Load by Migrating Birds Excrement,” *J. Korean Soc. Mar. Environ. Eng.*, **13**(4), 223~233 (2010).
 18. Yim, S. B., Kim, S. K., Seo, Y. S. and Park, S. H., “A study on Convergency of Tunnel Displacement using Control Chart Method,” *J. Eng. Geol.*, **17**(2), 197~204(2007).
 19. Huang, W., Mukherjee, D. and Chen, S., “Assessment of Hurricane Ivan impact on chlorophyll-a in Pensacola Bay by MODIS 250 m remote sensing,” *Mar. Pollut. Bullet.*, **62**, 490~498(2011).
 20. Rho, H. J., “Theory and Practice of Multivariate analysis by Excel and SPSS,” pp. 244~261(2005).
 21. Seo, J. K., Lee, H. K. and Chung, I. K., “The Community Dynamics of Microbial Food Web during Algal Bloom by *Stephanodiscus* spp. in Downstream of Nakdong River,” *Korean J. Environ. Biol.*, **28**(3), 172~178(2010).
 22. Cho, K. J. and Shin, J. K., “Persistent Blooms of Diatoms *Stephanodiscus Hantzschii* F. Tenuis and *S. Parvus* in the Nakdong River,” *Korean J. Phycol.*, **10**, 91~96(1995).
 23. Linda, E., Graham, James M. G. and Lee, W. W., *ALGAE* 2nd edition(2009).
 24. Zheng, B. H., Cao, C. J., Qin, Y. W. and Huang, M. S., “Analysis of nitrogen distribution characters and their sources of the major input rivers of three gorges reservoir,” *Environ. Sci.*, **29**(1), 1~6(2008).
 25. Hein, M., “Inorganic carbon limitation of photosynthesis in lake phytoplankton,” *Fresh Water Biol.*, **37**, 545~552(1997).
 26. Guoqi, H., Zhimin, M. and Nancy, C., “Hurricane Igor impacts on the stratification and phytoplankton bloom over the Grand Banks,” *J. Mar. Systems*, **100-101**, 19~25(2012).