

수문기상예측자료를 활용한 대청호 Chl-a 3개월 선행예측연구

곽재원[†]

환경부 한강홍수통제소

A Study on the 3-month Prior Prediction of Chl-a Concentration in the Daechong Lake using Hydrometeorological Forecasting Data

Kwak, Jaewon[†]

Han River Flood Control Office, Ministry of Environment, Seoul, Republic of Korea

(Received : 01 April 2021, Revised : 21 April 2021, Accepted : 21 April 2021)

요약

최근 반복되고 있는 녹조는 수질관리에 가장 큰 과제로서 대두되고 있다. 현재 환경부에서는 7일 단위의 선행수질예측을 통한 수질예보를 수행하고 있으나, 선제적인 조치를 위해서 좀 더 장기간의 수질예측이 필요한 시점이다. 이에 본 연구에서는 수질예보의 보완자료로서 대청호의 Chl-a 농도를 3개월 선행예측하기 위한 방법론을 제안하고 그 적용성을 검토하고자 한다. 이를 위하여 대청호의 수질자동측정망 자료와 ECMWF의 수문기상예측자료를 수집하였으며 각 시계열 자료의 특성을 분석하였다. 대청호의 Chl-a 농도와 상관 및 웨이블릿 분석을 바탕으로 수문기상입력인자를 결정하고 지연시간을 가지는 NARX모형을 이용하여 대청호의 Chl-a에 대한 3개월 선행예측 모형을 구축하였으며, 결과에 대한 비교분석을 통하여 모형의 적용성을 제시하였다.

핵심용어 : 계절예측, Chl-a 모델링, NARX 모형, 웨이블릿 분석

Abstract

In recently, the green algae bloom is one of the most severe challenges. The seven days prior prediction is in operation to issues the water quality warning, but it also needs a longer time of prediction to take preemptive measures. The objective of the study is to establish a method to conduct a 3-month prior prediction of Chl-a concentration in the Daechong Lake and tested its applicability as a supplementary of current water quality warning. The historical record of water quality in the Daechong Lake and seasonal forecasting of ECMWF were obtained, and its time-series characteristics were analyzed. The Chl-a forecasting model was established using a correlation between Chl-a concentration and meteorological factor and NARX model, and its efficiency was compared.

Key words : Seasonal forecasting, Chl-a modeling, NARX model, Wavelet analysis

[†]To whom correspondence should be addressed.
Han River Flood Control Office, Ministry of Environment
E-mail: firstsword@naver.com

• Kwak, Jaewon Han River Flood Control Office, Ministry of Environment / Researcher (firstsword@naver.com)

1. 서 론

지구온난화와 반복되는 가뭄으로 인하여 매년 여름마다 녹조가 발생하고 있으며, 그에 따른 수질안전 및 국민의 불안감도 나날이 증가하는 추세이다. 이에 환경부에서도 수질의 변화를 사전에 예측하고 수질악화가 예상되는 경우 선제적 조치를 하기 위한 7일 선행 수질예보를 2012년부터 수행하고 있으며, Chl-a는 주요한 항목의 하나로 간주되고 있다(Shin et al., 2017).

Chl-a의 증감은 기상 및 수문조건에 밀접한 영향을 받기 때문에 기온 및 강우 등 예측자료를 통해 조류의 발생을 사전에 예측이 가능하다. 과거 Chl-a 예측 관련 연구방법론은 크게 조류의 성장에 영향을 미치는 인자와의 상관성에 기반한 통계모형과 수치해석 기반의 결정론적 수치모델로 나누어져 왔다(Giri and Qiu, 2016; Lee et al., 2020). 통계모형을 통해 Chl-a 농도를 예측한 사례를 살펴보면, 주로 ANN(Artificial neural network)을 이용한 연구들이 많이 수행되었다(Coad et al. 2014). 초기에 Recknagel et al.,(1997)은 ANN을 이용하여 일본 Kasumigaura 호수에서 조류의 변동성을 파악하고 조류 발생에 대해 예측하였다. Jeong et al. (2003)은 evolutionary computation과 ANN을 이용하여 대한민국 낙동강의 조류 발생을 예측하였고, Velo-Suárez and Gutiérrez-Estrada(2007)은 조류 발생을 주단위로 예측하였다. 근래 들어서는 ANN(Asadollahfardi et al. 2018; Sinshaw et al. 2019; Goz et al. 2019), fuzzy logic (Huang et al. 2018), neuro fuzzy (Aghel et al. 2019; Kisi et al. 2019), multivariate adaptive regression spline (MARS) (Kisi and Parmar 2016; Najafzadeh and Ghaemi 2019), extreme learning machines (ELM) (Barzegar et al. 2018; Goz et al. 2019), random forest (Yajima and Derot 2018), support vector machines (LSSVM) (Fijani et al. 2019; Najafzadeh and Ghaemi 2019) 등 다양한 통계기법을 이용하여 수질예측의 적용성을 검증한 연구들이 수행되고 있으나, 수질인자 간의 상관관계가 내포하고 있는 불확실성에 큰 영향을 받고 있는 실정이다. 이에 반하여 수치모델링에 기반하는 방법은 여러 수질반응의 복잡성과 비선형성을 고려할 수 있기 때문에 다양한 모형이 제안되고 사용되어 왔다. 그 중에서도, 환경부에서는 하천·호소·하구에서의 수리·수질거동을 위해 세계적으로 널리 사용되는 EFDC모형을 기반으로 다양한 기능을 개선한 EFDC-NIER(Environmental Fluid Dynamic Code-National Institute of Environmental Research) 모형을 통하여 조류 발생을 예측하고 있으며(Shin et al., 2018), 해당 모형을 통하여 한강(Kim et al., 2014), 낙동강(Na et al., 2014), 금강(Kim et al., 2012) 등 주요 하천 수질예보에 활용되고 있다. 수치모델링에 기반하는 방법은 여러 가지 장점이 존재하고 있으나, 그에 반하여 많은 자료가 필요하고 구축이 어려우므로 녹조에 대해서 선제적으로 예측하고 대응하기 위한 장기간의 수질예측은 통계모형에 비해 더 어렵다는 단점이 있다(Giri and Qiu, 2016). 따라서,

현재 수질예보에서 시행되고 있는 7일 선행예측이 아니라 1개월 또는 3개월까지 장기간의 예측을 위한 보완적인 방법이 필요한 시점이라 하겠다.

이에 본 연구에서는 수질예보를 보완하는 3개월 선행예측을 위한 모형을 구축하고 그 적용성을 검토하고자 한다. 이를 위하여 대청호의 수질자동측정망 자료와 유럽연합의 ECMWF (European Centre for Medium-Range Weather Forecasts) 예측자료를 수집하였으며 각 시계열 자료의 특성을 확인하고, 대청호의 Chl-a 농도와 상관관계를 분석하였다. 교차웨이블릿분석을 바탕으로 수문기상입력자료를 결정하였으며, 지연시간을 가지는 NARX모형을 이용하여 대청호의 Chl-a에 대한 3개월 선행예측 모형을 구축하고 그 적용성을 검토하였다.

2. 이론적 배경

2.1 기후변동과 계절기후예측

기후변동(Climate variability)이란 긴 시간 동안의 평균값은 크게 벗어나지 않지만 약간의 변화를 보이는 자연적인 기후의 움직임에 말하며, 장기기간동안 나타나는 지속적인 평균값 및 특성의 변화인 기후변화(Climate change)와는 다르게 기후시스템의 내적인 과정으로 생기는 변동이다(Hegerl et al., 2007). 즉, 지구온난화 등으로 기인하는 인위적이고 영구적인 변동과는 달리 ENSO (El-Niño-Southern Oscillation)나 AMO (Atlantic Multidecadal Oscillation)와 같이 특정 주기를 가지며 자연적으로 발생하는 자연적 기후시스템의 변동을 가리키며, 이러한 기후변동을 예측하는 것이 기상예측의 주요한 목표 중 하나이다. 기상 및 기후예측모형은 예측시간 및 공간규모에 따라 일반적으로 단기예측, 중기예측, 장기예측으로 분류할 수 있다. 단기예측은 수 일 미만의 기상학적 변수에 대한 예측, 중기예측은 수 일에서 보름까지의 중규모 모형이나 기후모형의 예측결과에 의해 이루어진다. 마지막으로 장기예측(또는 계절예측)은 보름에서 수개월, 길게는 12개월까지의 예측을 말하며 주로 통계모형과 기후모형에 의해 이루어진다. 계절예측은 가뭄과 홍수, 폭염을 비롯한 기상이변으로 인한 재해를 대응하고 안정적인 수자원 확보와 관리를 위해 필수적인 정보이나, 아직 충분한 수준의 예측성을 확보하고 있지 못한 상태이며, 최근 지속적으로 관심과 연구가 집중되고 있는 분야 이다(Cho et al., 2016).

기후모형을 통한 계절예측 자료는 대표적으로 미국 NOAA (National Oceanic and Atmospheric Administration)의 NMME (the North American Multi-Model Ensemble) 자료와 유럽연합의 ECMWF 자료를 들 수 있다. 또한 호주 기상청(BOM, Bureau of Meteorology)에서도 POAMA (Prediction Ocean Atmosphere Model for Australia) 모형을 이용하여 9개월까지의 기후 예측 앙상블 자료를 매달 제공하고 있다(Cho et al., 2016). 또한 한국의 기상청 전지구 계절예측시스템(Global Seasonal forecast version 5,

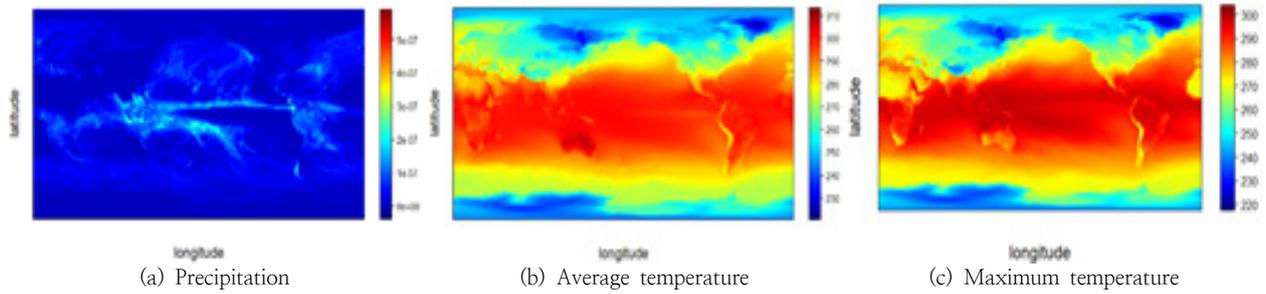


Fig. 1. ECMWF projection for global scale

GloSea5), APEC 기후센터(APCC, APEC Climate Center)에서 제공하는 다중모델 앙상블(MME, Multi Model Ensemble), K-water의 K-LPM(K-water Long-term Precipitation Model) 등이 장기기상전망 자료로 활용되고 있다(Jung, 2020). 본 연구에서는 Copernicus Climate Change Service (C3S)에서 제공하고 있는 ECMWF 계절예측 자료 중에서 가장 최근 버전이며, 매월 초에 향후 7개월까지의 일단위 계절예측 자료를 안정적으로 제공하고 있는 System5의 계절예측 일자료를 사용하였다. ECMWF 계절예측 자료는 예측력에 대한 많은 연구가 수행되었으며 타 기후모형의 계절예측 자료보다 예측력이 우수한 것으로 평가되었다 (Buizza and Chessa, 2002; Arpe et al, 2014; Lee et al, 2019; Jung, 2020). 수집한 예측 자료는 전지구 범위 자료이며 공간 해상도는 경위도 1° x 1° 격자 단위로 제공된다 (Fig. 1). 본 연구에서는 수집한 전지구 자료에서 대상지역인 대청호가 속한 금강유역을 포함하는 격자에 대해서 일단위의 강수량, 평균기온, 최대기온, 최소기온, 습도, 풍속, 태양복사량 및 증발량 자료를 추출하여 사용하였다. 대한민국의 경우 경위도 126.3° E~129.4° E, 34.4° N~38° N 범위에 위치하며 금강유역의 경우 중서부에 위치하고 있으므로 금강유역을 포함하는 127° E, 36° N 격자자료를 사용하였다.

2.2 검정통계량

Chl-a는 대부분의 광합성을 하는 조류에 포함되어 있어, 특정 수문기상인자와 명확한 관계를 보이지는 않으므로 (Willows, 2010), Chl-a 예측에 있어서 적합한 모형이나 인자를 선정하는 것은 난제이다. 따라서, 본 연구에서는 측정된 수문자료와 예측 인자로 사용되는 수문기상 인자의 특성을 파악하기 위하여 Ljung-Box Q 검정 (Ljung, 1978), 카오스 특성에 대한 BDS 통계 (Brock et al. 1991) 및 교차웨이블릿(Cross wavelet analysis(Bernstein, 1998) 분석을 적용하였다.

BDS(Brock-Dechert-Scheinkman) 검정은 시계열 자료가 무작위한 분포를 따른다는 가설을 상관적분을 토대로 검증하는 방법으로서, 무작위한 시계열 자료와 비선형 추계학적 시스템을 구분하는데 매우 유용하다(Brock et al., 1991). 시계열 자료가 무작위한 분포를 따른다는 가설 하에 $m > 1$ 인 경우의 BDS 검정은 다음과 같은 식(1) 과 식(2)으로 나타낼 수 있다.

$$BDS(m, N, r) = \frac{\sqrt{M}}{\sigma} [C(m, N, r) - C^m(1, M, r)] \quad (1)$$

$$C(m, N, r) = \frac{2}{M(M-1)} \sum_{1 < i < j < M}^M \theta(r - \|x_i - x_j\|) \quad (2)$$

$$\sigma^2 = 4m(m-1)C^{2(m-1)}(K - C^2) + K^m - C^{2m} + 2 \sum_{i=1}^{m-1} [C^{2i}(K^{m-i} - C^{2(m-i)}) - mC^{2(m-1)}(K - C^2)]$$

$$K = \frac{6 \sum_{1 < i < j < M}^M \theta(r - \|x_i - x_j\|)\theta(r - \|x_j - x_k\|)}{M(M-1)(M-2)}$$

여기서 $C(m, N, r)$ 은 반경 r 에 대한 상관적분이며, θ 는 단위계단함수, m, N, r 은 각각 위상공간차원, 좌표점의 개수, 위상반경을 나타낸다. 이를 시계열자료의 자기상관성을 검증하기 위하여 널리 사용되는 Ljung-Box 검정도 동시에 고려하면, 효율적으로 대상 시계열의 특성을 구분할 수 있다 (Brockwell et al., 2016).

웨이블릿분석(변환)은 특정한 시계열을 시간-주파수 공간으로 확장하는 방법으로 정의될 수 있으며 (Jevrejeva, 2003), 특히 국부적이거나 간헐적 주기성을 찾는 데 유용하게 사용될 수 있다 (Grinsted, 2004). 연속 웨이블릿 변환(CWT)는 식 (3)과 같이 나타낼 수 있다.

$$W_X(s, t) = X(t) \times \rho_s(t) \quad (3)$$

$$\rho_s(t) = \pi^{-\frac{1}{4}} \times e^{i\omega t} \times e^{-\frac{t^2}{2}}$$

여기서 $\rho_s(t)$ 는 스케일(s)를 가지는 Morlet 웨이블릿 (Lin, 2000)이며, 웨이블릿의 파워는 $|W_n^X(s)|^2$ 로 정의할 수 있다. 따라서 시계열 $X(t)$ 와 $Y(t)$ 사이의 교차웨이블릿은 $W_{XY}(s, t) = W_X(s, t) \cdot W_Y^*(s, t)$ 과 같이 나타낼 수 있으며, *는 공액복소수를 가리킨다. 이때 W_{XY} 의 복소편각은 시계열 $X(t)$ 와 $Y(t)$ 의 동일한 시간-빈도에서의 위상차로 해석된다.

2.3 NARX 모형

NARX 모형은 인공신경망의 한 종류로서 지연시간을 가지는 비선형 자기회귀모형을 포함하는 형태이다. 비선형 시계열 특성 모의에 적합한 특성을 보이며, 시계열 예측시 비선형성이 손실되는 문제에 대해서 자유로운 것으로 알려져 있다

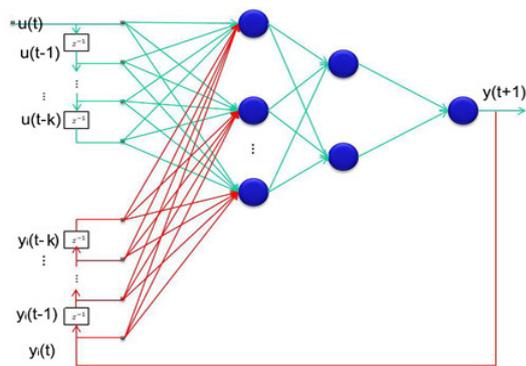


Fig. 2. NARX scheme

(Haykin, 1999). NARX 모형의 구조는 Fig. 2를 통하여 설명할 수 있다.

$y_t(t+1)$ 과 $y_t(t)$ 는 지연시간($i=1,2,\dots,k$)을 고려한 회귀신경망의 결과이며, t 는 시계열의 시간을 나타낸다. 따라서, NARX모형은 이전 시간의 결과를 고려하는 재순환적 동적 신경망이라 할 수 있다. 신경망의 구축에 핵심적인 역할을 수행하는 역전파 알고리즘은 수문기상 전반의 비선형 문제에 대해서 좋은 결과를 나타내는 것으로 알려진 Levenberg-Marquardt-QNBP 알고리즘을 적용하였다 (Battiti, 1989). 본 연구에서의 신경망은 1~5주의 지연시간 (k)를 갖는 피드백 층을 포함하여 총 16개의 은닉층을 가지는 것으로 설정 되었으며, 학습시에는 관측값을 이용하는 Open-loop, 실제 3개월 예측 시에는 이전 예측치를 이용하는 Close-loop로 구성되었다.

3. 자료 구축

3.1 대상유역 및 자료

대청호는 금강수계 수자원의 다목적 개발을 위해 1980년

대청댐이 완공되며 생겨난 대한민국에서 2번째로 큰 다목적 인공호수이다(Fig. 3 참조). 지리적으로는 대전시에 인접해 있으며 대전을 비롯한 충청권 대부분 지역에 연간 약 16억 m^3 의 용수를 공급하는 수원으로 상수원보호구역으로 지정되어 수질보전을 위해 각종 시설의 입지 및 토지이용이 규제되고 있다. 대청호의 유역면적은 4,134 km^2 로 저수면적에 대한 유역면적의 비가 56.8에 달하여 상대적인 오염의 가능성이 크며, 총저수량에 대한 유효저수량의 비율이 53.0%로 용수 이용율이 낮은 호수이다(Ministry of Environment, 2021a, 2021b). 특히 대청호는 지형적 특성으로 인해 체류시간이 길어 조류가 발생하기 유리한 조건을 갖추고 있어 최근 여름철 폭염에 따른 녹조문제가 대두되고 있다.

분석을 위한 기상자료는 ECMWF 자료의 가용기간인 2012년부터 2020년까지에 대해 평균, 최고, 최소기온 및 강수량, 상대습도, 평균풍속, 태양복사량, 증발산량의 일자료를 수집하였으며, 수질자료는 환경부 수질 자동측정망 중 대청호 지점에 대하여 기상자료와 동일한 기간의 일단위 Chl-a 자료를 수집하였다. 수집된 자료는 수질예측이 매주 시행되는 점을 고려하여 주 단위 자료로 변환하였다. 그 외에 연구를 위하여 필요한 유역의 특성이나 기상학적 특성의 경우에는 국가수자원종합정보시스템(Ministry of Environment, 2021a), MyWater(K-Water, 2021), 물환경정보시스템(Ministry of Environment, 2021b)을 통하여 수집하였다.

3.2 수질예측을 위한 모형·인자선정

비선형 시계열자료에 대해서 자료의 특성을 반영할 수 있는 적합한 모형과 인자를 선정하는 것은 중요한 문제이다 (Judd and Mees, 1995). Chl-a에 대해서도 기존의 연구에서 비선형 시계열임이 제시되었으나 (Song and Han, 2017; Lee et al., 2014), 대청호의 Chl-a 및 수문기상 시계열에 대한 확인을 위하여 BDS 통계 및 LBQ 검정을 시

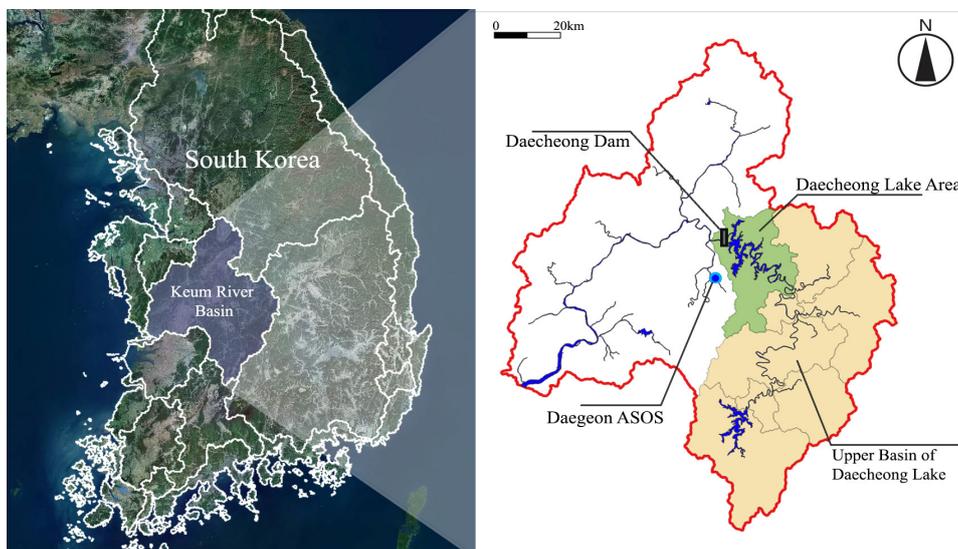


Fig. 3. Study area and basin

행하고 그 결과를 Table 1에 나타내었다. 검증결과에 따르면, Chl-a 및 최고, 평균, 최저기온과 증발량은 비선형시계열 또는 카오스시계열인 것으로 나타났으며, 강수량, 습도, 풍속 및 태양복사량은 무작위시계열인 것으로 나타났다. 다만, 태양복사량의 경우에는 일부 추계학적 특성도 보이고 있다. 따라서, 대청호의 Chl-a 시계열은 비선형 시계열 특성을 보이고 있으므로 예측 및 모의를 위해서는 해당 특성을 반영할 수 있는 모형이 필요할 것으로 판단된다.

또한, 많은 연구에서 Chl-a를 비롯한 수질항목은 수문기상인자와 높은 연관성이 있음을 지적하였다. 대표적으로 Song and Han (2017)은 기온, 강수량, 일사량, 유량, DO, BOD, TN, TP가 유의미한 관계가 있음을 나타내었으며,

특히, 단기적으로는 Chl-a 자체의 영향, 장기적으로는 TP가 가장 큰 영향을 미치는 것으로 제시하였다. 다만, Chl-a의 경우에는 시간적인 경향성과 상관관계의 유무에 대해서 상반된 결과를 제시하는 연구가 존재하고 있다 (Park, 1998; Kwon et al., 2014; Oh and Kim, 1995; Park and Kim, 2013). 이는 Chl-a 시계열이 전체적인 계절성을 보이지않고, 국부적인 계절성과 상관성을 가지고 있는 것이 주요한 원인으로 판단된다 (Li et al., 2019). 따라서, 이러한 경향을 고려하여 인자를 선정하기 위하여 교차웨이브릿 분석을 시행하였다 (Fig. 4). 크게 평균기온, 최고기온, 강수량 등이 높은 국부 상관성을 보이는 것으로 나타났으며, 최저기온, 태양복사량, 증발량 등은 낮은 국부 상관성, 습도와

Table 1. Test statistics for each time series

Test Stat.	Item									95% C.I.
	Chl-a	Ave. temp.	Min. temp.	Max. temp.	Prcp.	Humid.	Wind speed	Solar Radiation	Evapor.	
BDS(1)	11.58	4.06	2.78	7.76	0.39	0.05	0.49	1.96	4.57	[+-1.96]
BDS(2)	10.81	3.96	2.58	7.02	0.43	0.07	0.47	2.24	5.82	[+-1.96]
BDS(3)	9.79	3.66	2.35	6.32	0.49	0.96	0.65	1.94	7.57	[+-1.96]
BDS(4)	8.97	3.39	2.20	5.77	0.54	0.70	0.39	1.80	10.89	[+-1.96]
BDS(5)	8.68	3.17	2.13	5.33	0.58	0.55	0.92	1.59	17.09	[+-1.96]
LBQ(1)	23.62	32.77	32.56	32.21	2.82	1.23	2.3	1.99	15.35	3.8
LBQ(2)	37.81	63.10	62.67	61.83	3.22	2.09	3.9	2.41	27.11	6.0
LBQ(3)	46.72	90.35	89.72	88.32	3.23	2.84	5.8	3.08	37.42	7.8
LBQ(4)	52.70	113.86	113.11	111.06	3.23	3.32	7.3	3.16	45.04	9.5
LBQ(5)	57.06	133.26	132.45	129.66	3.43	3.74	8.2	3.17	50.52	11.1
Check	Nonlinear stochastic or chaotic time-series				Random series				Nonlinear time-series	

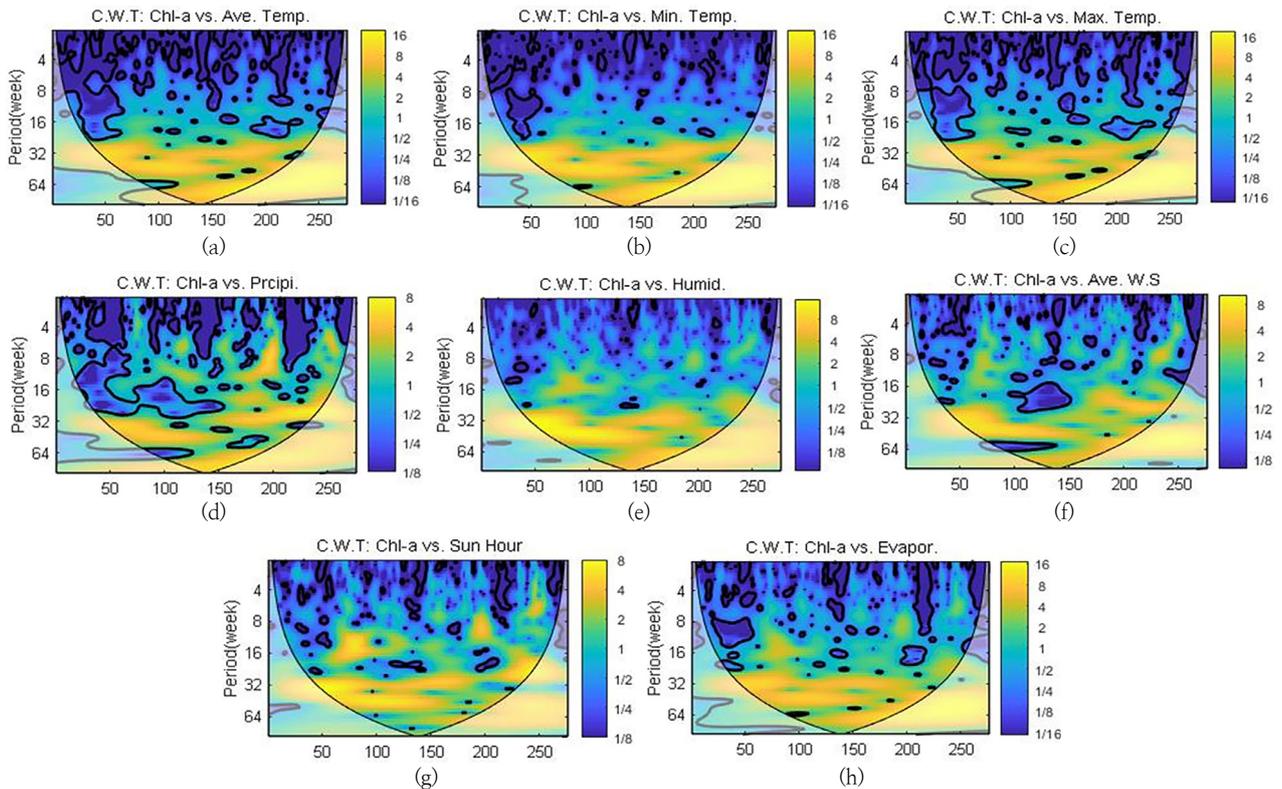


Fig. 4. Cross-wavelet diagram for Chl-a and meteorological factor

평균풍속은 미미한 상관성을 보여주었다. 특징적으로 모든 상관관계에서 겨울철은 상관성이 없는 것으로 나타났으며, 그 외의 기간에 주기적 상관성을 나타냈다. 이는 Chl-a가 봄-가을 사이의 수문기상인자에 영향을 받는다는 것을 의미한다. 상기의 검증결과를 고려하면, Chl-a을 예측하기 위해서 평균·최고기온, 강수량을 입력인자로 비선형 시계열특성을 고려하는 모형을 이용하는 것이 적합할 것으로 판단된다. 특히, Song and Han (2017)의 연구에서 제시된 바와 같이 Chl-a 자체영향이 큰 점을 고려하면, 지연시간이 있는 피드백을 포함하는 NARX 모형이 예측에 적합할 것으로 판단된다. 따라서, 본 연구에서는 NARX 모형을 적용하였고, 모형의 주요 입력인자는 평균·최고기온 및 강수량이다.

4. 모형구축 및 적용결과

4.1 모형구축 및 보정

대청호의 Chl-a를 예측하기 위하여 ECMWF의 3개월 계절예측자료와 NARX모형을 이용하여 모형 구축 및 보정을 수행하였다. 2013년부터 2020년까지의 Chl-a 및 ECMWF 계절자료를 수집하였으며, 2013년부터 2018년까지의 기간을 모형보정기간, 2019년을 모형검증기간으로 구축하였다 (2020년 제외).

Chl-a 예측을 위한 NARX 모형의 경우에는 Trial and error 방법을 이용하여 총 16개의 은닉층을 가지는 것으로 구성하였으며, 역전파알고리즘은 비선형 문제에 대해서 좋은 결과를 나타내는 것으로 알려진 Levenberg-Marquardt-QNBP 알고리즘을 적용하였다. 모형 지연시간은 Song and Han (2017)의 연구에 따라서 1~5주의 지연시간을 가지는 것으로 설정하였다. 모형의 입력자료는 이전 절에서의 분석결과에 따라서, ECMWF 계절예측자료(hindcast) 중에서 평균기온, 최고기온, 강수량 자료를 이용하였다. 기존의 연구를 고려하면 상기의 기온과 강수량 이외에도 하천 유량과 영양염류 등이 영향을 미치는 것으로 제시된 바 있다 (Cho et al., 2020). 그러나, 하천유량의 경우에는 강우에 따른 종속적인 변수이며 유역유출에 따른 불확실성을 내포하고 있으므로 중첩된 불확실성으로 인하여 장기간 예측의 신뢰성 확보는 어려운 실정이다 (Lee et al., 2019). 따라서, 예측유출량은 강수량에 대한 종속변수임을 고려하여 제외하였다. TN이나 TP 등의 영양염류유입도 조류발생과 상관관계가 있으나 농축산연건이나 강우초기 비점오염원 유입, 수질오염사고, 퇴적물에서의 용출 등 수많은 영향인자가 복합적으로 작용하므로 그 거동을 예측하기 어렵고 (Han et al., 2008) 강수량에 따라서 유입이 증가하는 경우가 많으므로, 역시 종속적인 변수로 간주하여 제외하였다. 이러한 변수의 제외는 필연적으로 예측력에 저하를 가져오게 되나, 기존 연구사례에서 모형의 불확실성을 고려하여 오히려 적은 입력변수를 선정하였다는 점 (Kim and Kim, 1999)을 감안

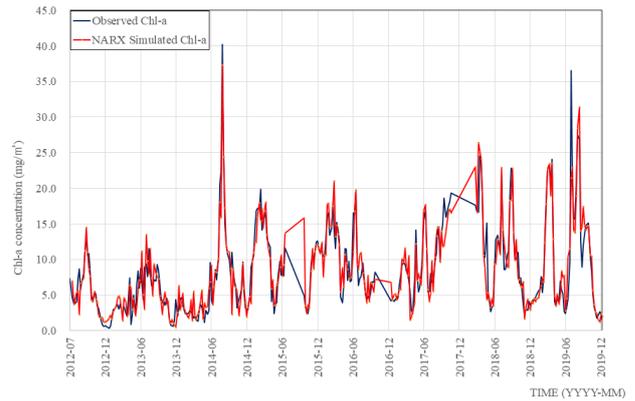


Fig. 5. Calibration result of NARX model for Chl-a

하면 제외 가능할 것으로 판단된다. 또한, 대청호의 Chl-a 자료에는 3개의 결측기간 ('15, '16, '17년)이 있었으며 해당 기간은 분리하여 모의하였다. 상기의 조건을 이용하여 모형을 보정하였으며, 그 결과는 Fig. 5와 같다. Fig 5의 보정결과를 보면, 전체적으로 대청호의 Chl-a를 잘 모의하는 것으로 보인다. 또한, 모형의 효율성을 판단하기 위하여 널리 활용되는 RMSE와 R^2 (Hyndman and Khandalar, 2006)의 경우 각각 2.56과 0.82의 우수한 결과를 나타내었다. 따라서, 수문기상인자를 활용해서 Chl-a의 모의가 가능할 것으로 판단된다.

4.2 검증 및 토의

구축된 모형의 효율을 검증하기 위하여, 이전 절에서 보정완료된 Chl-a 예측모형과 ECMWF의 계절예측자료를 이용하여 2019년도 대청호의 Chl-a 농도에 대한 3개월 선형예측을 수행하였다. 연구가 수행되었을 시점에서 2020년도의 비검증된 Chl-a 자료도 취득할 수 있으나 2020년 자료의 Chl-a 거동 자체가 예년에 비하여 상이하여 제외하였다 (Fig. 6 참조).

이는 ECMWF 계절예측자료의 2020년도 불확실성이 높은 것이 원인으로 보인다. 실제로 2020년도 ECMWF 계절예측에서의 연강수량은 1,200mm 선인데 비하여, 실제 강수량은 1,500mm이상으로 300mm 이상의 차이가 발생하였으며,

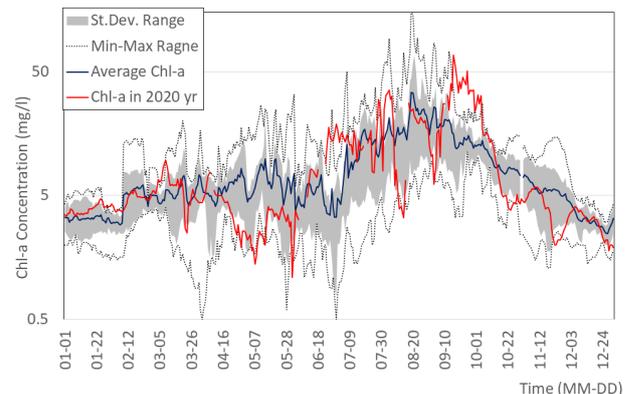


Fig. 6. Chl-a time series of 2012 to 2019 versus 2020 year

강수의 시점이나 집중 정도 차이가 커 적용하기 어렵다. 이는 2020년도의 여름철 강수량이 이례적인 수준인 것이 원인으로 판단되며, 현재 기후예측시스템의 한계에 따른 것으로 보인다. 실제로 ECMWF를 비롯한 계절예측관련 연구에서도 기온의 경우에는 충분히 계절 규모에서 활용할 수 있는 수준이나, 강수의 경우에는 일부에 한하여 신뢰할 수 있는 수준인 것으로 제시하기도 하였다 (Hyun et al., 2020). 본 연구에서와 같이 계절 단위의 수질예측 정확도는 기상예측에 크게 의존하고 있으므로 예측정확도 향상과 수질예측에 적합한 자료를 위한 협력이 지속적으로 필요할 것이다.

모형의 검증은 2019년 1월부터 12월까지 매주 월요일에 ECMWF 계절예측자료(Forecast)를 이용하여 3개월의 수질예보를 수행하는 것을 가정하였으며, 52주 동안의 예측결과를 비교하는 방식으로 수행하였다. Fig. 7(a)는 2019년 매주 대청호의 Chl-a 농도를 3개월 선행 예측하였다고 가정하였을 경우의 각각의 선행예측 곡선과 관측값을 나타낸 것이며, Fig. 7(b)는 각각의 선행예측 곡선에 대해서 통계적인 범위를 나타낸 것이다. ECMWF의 계절예측자료를 활용하기 때문에 모형 자체의 불확실성과 예측자료의 불확실성을 모두 내포하고 있다고 할 수 있다. 따라서, 이러한 불확실성을 구분하기 위하여 동일한 모형에 대하여 2019년의 ECMWF 관측값(Hindcast)를 적용한 모의결과도 동시에 나타내었다. Fig. 7의 검증결과를 보면 전체적으로 2019년의 Chl-a 농도를 설득력 있게 예측한 것으로 판단된다. 관측자료를 이용한 NARX모형의 RMSE와 R^2 값은 각각 3.48과 0.81을 나타내었으며, 이는 기존의 모형보정기간과 유사한 결과이다. 계절예측자료를 사용한 3개월 선행예측 평균 RMSE와 R^2 는 각각 4.1과 0.71로 다소 낮게 나타났으나 적용성은 충분한 수준으로 판단된다. 따라서, 체계적 기반의 Chl-a 계절예측을 적합하게 수행할 수 있을 것으로 판단된다.

세부적으로 보면 1, 2, 12월과 같이 녹조가 거의 발생하지 않는 기간은 유사하게 모의하고 있으나, 3~4월 이후에는 Chl-a 농도 상승 및 유지기간에 다소 큰 불확실성을 나타내었으며, 7월말 Chl-a농도의 상승경향은 유사하게 모의하였으나 급격한 상승을 예측하지 못하였다. 또한, 9월 이후에 모형의 불확실성이 급격하게 커지는 모습을 보여주었다. 3월부터 나타나는 Chl-a의 상승경향은 관측자료를 이용한 모의에서는 보이지 않았던 것을 고려하면 계절예측자료로 인한 불확실성으로 판단된다. 그러나, 7월의 급격한 Chl-a 상승은 계절예측 및 관측값으로도 모두 모의하지 못하였으므로 모형 자체의 불확실성으로 판단된다. 이 경우, 추정가능한 원인으로는 7월 강수로 인하여 급격하게 유입된 영양염류의 영향을 의심해 볼 수 있으나, 차후 추가적인 연구를 통하여 개선하여야 할 것으로 판단된다. 또한, 9월~11월 사이에 모형의 불확실성이 급격하게 증가하는 경향을 보이는데 역시 계절예측과 관련된 것으로 보이며, 한반도의 기상학적 특성으로 인하여 해당 기간의 계절예측 불확실성이 높은 것이 주요 원인으로 판단된다. 실제로 최근 연구에서 중위도에서 내륙으로 수증기를 운반하는 대기천(Atmospheric rivers)이 한반도 연 강우량 57%에 영향을 주고, 특히 여름철 강우량에 큰 영향을 주는 것으로 제시되었으며 (Han et al., 2020), 태풍으로 인한 연강수량의 영향도 18%인 것으로 제시되었다 (Moon, 2019). 따라서, 향후 지속적인 계절예측 연구를 통하여 해당 불확실성을 개선해야 할 것으로 판단된다. 이를 종합적으로 판단하여 보면, 계절예측자료와 NARX모형을 이용한 대청호 Chl-a 농도에 대한 3개월 선행예측은 적용성이 있는 것으로 판단된다. 다만, 여름철에 국지성 호우나 태풍 등으로 인한 불확실성이 존재하고 있으므로 이를 충분히 고려하여야 할 것으로 보인다.

본 연구의 보완점은 예측자료의 불확실성이라고 할 수 있다. 연구에서 제안된 대청호 Chl-a의 예측모형은 수문기상인자만을 사용하고 있다. 이는 영양염류의 유입을 예측하기 어렵기 때문이나, 평년수준의 영양염류 유입에서 벗어나는 유입이 발생할 경우 설득력 있는 예측이 어렵다. 이는 모형의 입력인자로 사용되는 계절예측자료에도 동일하게 적용된다. 입력자료로 ECMWF의 계절예측자료를 사용하고 있으므로, 태풍과 집중호우로 인하여 발생하는 불확실성을 지속적으로 감안하여야 하는 단점이 있다. 그러나, 본 연구는 조류대응에 핵심이 되는 Chl-a 예측에 있어서 기존의 수치모형이 아닌 3개월까지의 선행예측을 간편하게 수행할 수 있다는 점에서 그 의미가 있으며, 향후 영양염류 유입과 계절예측의 정확성이 확보된다면 수질예보의 보완적인 자료로서 활용성을 확보할 수 있을 것이다.

5. 결 론

본 연구에서는 대청호 내의 Chl-a 농도를 3개월까지 선행예측 하기 위한 모형을 구축하고 그 적용성을 검토하였

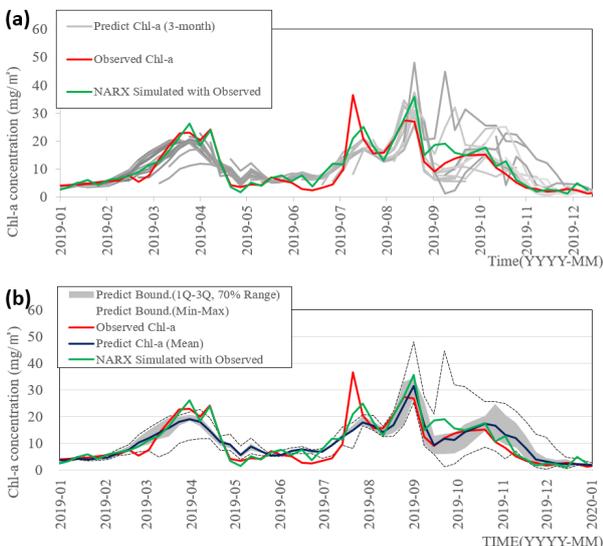


Fig. 7. Validation result of NARX model for Chl-a

다. 이를 위하여 2012년부터 2019년까지의 대청호의 일 단 위 수질자동측정망과 ECMWF의 수문기상예측(계절예측) 자료를 수집하였다. 이를 주 단위로 변환하여 수문기상자료와 Chl-a 농도 사이의 상관관계를 분석하였으며, BDS통계와 LBQ 검정을 이용하여 Chl-a를 비롯한 주요 수질인자가 비선형 시계열 특성을 가지고 있음을 확인하고, 예측을 위해서는 비선형 시계열을 고려할 수 있는 모형을 사용하여야 함을 제시하였다. 또한, Chl-a 시계열 및 수문기상인자의 교차웨이블릿 분석을 통하여 선행예측을 위해서 평균기온, 최대기온 및 강수량을 입력인자로 선정하고 NARX 모형을 이용하여 2012~2018년 기간동안 모형을 보정하고, 2019년 자료를 통하여 그 적용성을 검증하였다. 모형의 효율을 판단할 때 주로 사용하는 RMSE와 R^2 는 보정기간에 2.56과 0.82, 검증기간에는 4.1와 0.71을 보여주어 해당 모형이 대청호의 Chl-a 농도를 3개월 선행예측하기 위한 적용성이 있는 것으로 나타났다. 그러나, 태풍을 비롯한 집중호우의 영향으로 9~11월에 높은 예측 불확실성을 보이며, 모형 자체의 예측신뢰도도 기상예측자료 정확성에 크게 의존하고 있으므로 신뢰성있는 예측자료의 확보에도 노력을 기울여야 할 것으로 보인다. 따라서, 녹조가 주로 발생하는 여름철에 불확실성이 존재하고 있어 예측결과에 대한 종합적인 검토를 거쳐야 할 것으로 판단된다. 그러나, 현행 수질예보에 사용되는 수치기반 모형보다 간편하게 3개월까지 수질 선행예측이 가능하다는 점에서 보완자료로서 활용은 가능할 것으로 판단되며, 본 연구를 통해서 제시된 방법론 및 모형이 보완자료로서 활용할 수 있을 것을 기대한다.

References

- Aghel, B., Rezaei, A., and Mohadesi, M. (2019). Modeling and prediction of water quality parameters using a hybrid particle swarm optimization-neural fuzzy approach. *International Journal of Environmental Science and Technology*, Vol. 16, No. 8, pp. 4823-4832.
- Asadollahfardi, G., Zangooi, H., Asadi, M., Tayebi Jebeli, M., Meshkat-Dini, M., and Roohani, N. (2018). Comparison of Box-Jenkins time series and ANN in predicting total dissolved solid at the Zāyandē-Rūd River, Iran. *Journal of Water Supply: Research and Technology-Aqua*, Vol. 67, No. 7, pp. 673-684.
- Arpe, K., Leroy, S. A. G., Wetterhall, F., Khan, V., Hagemann, S., and Lahijani, H. (2014). Prediction of the Caspian Sea level using ECMWF seasonal forecasts and reanalysis. *Theoretical and applied climatology*, Vol. 117, No. 1, pp. 41-60.
- Battiti, R., 1994. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks*, Vol. 5, No. 4, pp. 537-550. <http://dx.doi.org/10.1109/72.298224>.
- Barzegar, R., Moghaddam, A. A., Adamowski, J., and Ozga-Zielinski, B. (2018). Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model. *Stochastic environmental research and risk assessment*, Vol. 32, No. 3, pp. 799-813.
- Bernstein, J. (1998). *Seasonality: Systems, Strategies and Signals*, John Wiley & Sons, NY, USA.
- Brock, W. A., Hsieh, D. A., LeBaron, B. D., and Brock, W. E. (1991). *Nonlinear dynamics, chaos, and instability: statistical theory and economic evidence*. MIT Press, ME, USA.
- Brockwell, P. J., and Davis, R. A. (2016). *Introduction to time series and forecasting*. Springer, NY: USA.
- Buizza, R., and Chessa, P. (2002). Prediction of the US storm of 24-26 January 2000 with the ECMWF Ensemble Prediction System. *Monthly weather review*, Vol. 130, No. 6, pp. 1531-1551.
- Cho, H., Huong, N. T., Moon, J., and Kwon, H. H. (2020). Prediction of cyanobacteria population based on Poisson regression based on hydro-meteorological condition. *In Proceedings of the Korea Water Resources Association Conference*, Jeju, Korea, Vol. 1, pp. 208.
- Cho, J., Jung, I. W., Kim, C. G., and Kim, T. G. (2016). One-month lead dam inflow forecast using climate indices based on tele-connection. *Journal of Korea Water Resources Association*, Vol. 49, No. 5, pp. 361-372.
- Coad, P., Cathers, B., Ball, J. E., and Kadluczka, R. (2014). Proactive management of estuarine algal blooms using an automated monitoring buoy coupled with an artificial neural network. *Environmental modelling & software*, Vol. 61, pp. 393-409.
- Fijani, E., Barzegar, R., Deo, R., Tziritis, E., and Skordas, K. (2019). Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters. *Science of the total environment*, Vol. 648, pp. 839-853.
- Grinsted, A., Moore, J.C., and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlin. Process. Geophys.* Vol. 11, pp. 561-566.
- Giri, S., and Qiu, Z. (2016). Understanding the relationship of land uses and water quality in twenty first century: a review. *Journal of environmental management*, Vol. 173, pp. 41-48.
- Go, E., Yu, M., and Karadurmus, E. (2019). Total organic carbon prediction with artificial intelligence techniques. *In Computer aided chemical engineering*, Vol. 46, pp. 889-894.
- Han, J. S., Kim, C. S., and Han, G. S. (2008). Pollution

- Control & Remediation of Contaminated Groundwater. *Geoenvironmental Engineering*, Vol. 9, No. 3, pp. 5–21.
- Hegerl, G. C., Crowley, T. J., Allen, M., Hyde, W. T., Pollack, H. N., Smerdon, J., and Zorita, E. (2007). Detection of human influence on a new, validated 1500-year temperature reconstruction. *Journal of Climate*, Vol. 20, No. 4, pp. 650–666.
- Huang, M., Tian, D., Liu, H., Zhang, C., Yi, X., Cai, J., and Ying, G. (2018). A hybrid fuzzy wavelet neural network model with self-adapted fuzzy-means clustering and genetic algorithm for water quality prediction in rivers. *Complexity*, Vol. 2018, <https://doi.org/10.1155/2018/8241342>.
- Han, H., Choi, C., Moon, H., Jung, J., Lee, C., and Kim, H. S. (2020). Hydrological impact of Atmospheric River landfall on the Korean Peninsula. *Journal of Korea Water Resources Association*, Vol. 53, No. 11, pp. 1039–1047.
- Hyndman, R.J. and Khandakar, Y., (2006). *Automatic time series for forecasting: the forecast package for R (No. 6/07)*. Monash University, Melbourne: Department of Econometrics and Business Statistics, Melbourne, Australia.
- Hyun, Y. K., Park, J., Lee, J., Lim, S., Heo, S. I., Ham, H., and Kim, Y. (2020). Reliability Assessment of Temperature and Precipitation Seasonal Probability in Current Climate Prediction Systems. *Atmosphere*, Vol. 30, No. 2, pp. 141–154.
- Jeong, K. S., Kim, D. K., Whigham, P., and Joo, G. J. (2003). Modelling Microcystis aeruginosa bloom dynamics in the Nakdong River by means of evolutionary computation and statistical approach. *Ecological Modelling*, Vol. 161, No. 1–2, pp. 67–78.
- Jevrejeva, S.; Moore, J.C., and Grinsted, A. (2003). Influence of the Arctic Oscillation and El Niño–Southern Oscillation (ENSO) on ice conditions in the Baltic Sea: The wavelet approach. *Journal of Geophysical Research: Atmospheres*. Vol. 108, pp. D21.
- Judd, K., and Mees, A. (1995). On selecting models for nonlinear time series. *Physica D: Nonlinear Phenomena*, Vol. 82, No. 4, pp. 426–444.
- Jung, J. W. (2020). *Seasonal climate prediction reflecting climate variability and flexible operation plan of a multi-purpose dam in flood season*, Ph.D Dissertations, INHA University, Incheon, Korea.
- Kisi, O., and Parmar, K. S. (2016). Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. *Journal of Hydrology*, Vol. 534, pp. 104–112.
- Kisi, O., Azad, A., Kashi, H., Saeedian, A., Hashemi, S. A. A., and Ghorbani, S. (2019). Modeling groundwater quality parameters using hybrid neuro-fuzzy methods. *Water resources management*, Vol. 33, No. 2, pp. 847–861.
- Kim, B. D., and Kim, S. O. (1999). Model Selection Accounting for Model Uncertainty. *Journal of Korean Marketing Association*, Vol. 14, No. 1, pp. 93–107.
- Kim, D. W., Chung, E. G., and Kim, K. H. (2012). Application of Hydrodynamic and Water Quality Model EFDC in Geum River, *Proceedings of the 2012 Spring Co-Conference of Korean Society on Water Environment and Korean Society of Water and Wastewater, Korean Society on Water Environment and Korean Society of Water and Wastewater*, Ilsan, Korea, Vol. 1, pp. 782–783.
- Kim, K., Park, M., Min, J-H., Ryu, I., Kang, M-R., and Park, L.J. (2014). Simulation of algal bloom dynamics in a river with the ensemble Kalman filter, *Journal of Hydrology*, Vol. 519, pp. 2810–2821.
- Kwon, Y. S., Bae, M. J., Kim, J. S., Kim, Y. J., Kim, B. H., and Park, Y. S. (2014). Characterizing changes of water quality and relationships with environmental factors in the selected Korean reservoirs. *Korean Journal of Ecology and Environment*, Vol. 47, No. 3, pp. 146–159.
- Korea Meteorological Administration (2021). Open MET Data Portal, accessed 11 Feb. 2021, <<https://data.kma.go.kr/>>.
- K-Water (2021). K-Water Corporation, accessed 20 February 2021, <<https://www.mywater.or.kr/>>.
- Lee, T. H., and Kim, H. J., and Kim, M. (2014). Relationships Between the Chl-a and Influence Factors and Characteristics of Algae Occurrence in Banwol Reservoir. *Proceedings of the 2014 KSAE Annual Conference*, Byeonsan, Korea. Vol. 2014, pp. 201.
- Lee, S., Kim, G., Yoon, S., and An, H. (2019). Improvement of precipitation forecasting skill of ECMWF data using multi-layer perceptron technique. *Journal of Korea Water Resources Association*, Vol. 52, No. 7, pp. 475–482.
- Lee, H. Y., Jang, S. H., Lee, J. K., and Jo, J. W. (2019). Probabilistic Runoff Analysis using Ensemble Technoque with Localization Method. *In Proceedings of the Korea Water Resources Association Conference*, Yeosu, Korea, Vol. 1, pp. 207.
- Lee, S. M., Park, K. D., and Kim, I. K. (2020). Comparison of machine learning algorithms for Chl-a prediction in the middle of Nakdong River (focusing on water quality and quantity factors). *Journal of Korean Society of Water and Wastewater*, Vol. 34, No. 4, pp. 277–288.
- Lin, J., Qu, L. (2000). Feature extraction based on Morlet wavelet and its application for mechanical fault

- diagnosis. *J. of Sound Vib.* Vol. 234, pp. 135–148.
- Li, L., Jiang, P., Xu, H., Lin, G., Guo, D., & Wu, H. (2019). Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang River, China. *Environmental Science and Pollution Research*, Vol. 26, No. 19, pp. 19879–19896.
- Ministry of Environment (2021a). Water Resources Management Information System, accessed 17 Feb. 2021, <<http://www.wamis.go.kr/>>.
- Ministry of Environment (2021b). Water Environment Information System, accessed 21 Feb. 2021, <<http://water.nier.go.kr/>>.
- Mun, J. W. (2018). A Study on the Contribution of Typhoon to Annual Precipitation in Korea. *In Proceedings of KSCE 2018 Convention*, Kyoungju, Korea, Vol. 1, pp. 131–132.
- Na, E. H., Shin, C. M., Park, L. J., Kim, D. G., and Kim, K. H. (2014). A Study on the Operational Forecasting of the Nakdong River Flow with a Combined Watershed and Waterbody Model, *Journal of Korean Society on Water Environment*, Vol. 30, No. 1, pp. 16–24.
- Najafzadeh, M., and Ghaemi, A. (2019). Prediction of the five-day biochemical oxygen demand and chemical oxygen demand in natural streams using machine learning methods. *Environmental monitoring and assessment*, Vol. 191, No. 6, pp. 1–21.
- Oh, H. M. and Kim, D. H. (1995). Short-term prediction of the blue-green algal bloom in Daechung Reservoir. *Korean Journal of Limnology*, Vol. 28, No. 2, pp. 127–135.
- Park, S.Y. (1998). *Trend analysis of water qualities in major lakes of Korea with seasonal Mann-Kendall test*. Master thesis, Ewha Womans University, Seoul, Korea.
- Park, H. K., and Kim, E. K. (2013). Study on the modelling of algal dynamics in lake paldang using artificial neural networks. *Journal of Korean Society on Water Environment*, Vol. 29, No. 1, pp. 19–28.
- Recknagel, F., French, M., Harkonen, P., and Yabunaka, K. I. (1997). Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling*, Vol. 96, No. 1–3, pp. 11–28.
- Sinshaw, T. A., Surbeck, C. Q., Yasarer, H., and Najjar, Y. (2019). Artificial neural network for prediction of total nitrogen and phosphorus in US Lakes. *Journal of Environmental Engineering*, Vol. 145, No. 6, 04019032.
- Shin, C. M., Min, J. H., Park, S. Y., Choi, J., Park, J. H., Song, Y. S., and Kim, K. (2017). Operational water quality forecast for the Yeongsan river using EFDC model. *Journal of Korean Society on Water Environment*, Vol. 33, No. 2, pp. 219–229.
- Song, J. H., and Han, I. S. (2017). Prediction and Characterization Analysis of River Water Quality using Multivariate Time Series Models. *Journal of the Korean Society of Urban Environment*, Vol. 17, No. 3, pp. 243–253.
- Song, J. H. (2017). *Prediction and Characterization Analysis of River Water Quality using Multivariate Time Series Models*, Master thesis, University of Seoul, Seoul, Korea.
- Velo-Suárez, L., and Gutiérrez-Estrada, J. C. (2007). Artificial neural network approaches to one-step weekly prediction of *Dinophysis acuminata* blooms in Huelva (Western Andalucía, Spain). *Harmful Algae*, Vol. 6, No. 3, pp. 361–371.
- Willows, R. D. (2003). Biosynthesis of chlorophylls from protoporphyrin IX. *Natural product reports*, Vol. 20, No. 3, pp. 327–341.
- Yajima, H., and Derot, J. (2018). Application of the Random Forest model for chlorophyll-a forecasts in fresh and brackish water bodies in Japan, using multivariate long-term databases. *Journal of Hydroinformatics*, Vol. 20, No. 1, pp. 206–220.