Journal of Hydrology 529 (2015) 805–815

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Stressor–response modeling using the 2D water quality model and regression trees to predict *chlorophyll-a* in a reservoir system

Yongeun Park [a], Yakov A. Pachepsky [a], Kyung Hwa Cho [b], Dong Jin Jeon [c], Joon Ha Kim [c,*]

[a] USDA-ARS, Environmental Microbial and Food Safety Laboratory, 10300 Baltimore Ave., Building 173, BARC-EAST, Beltsville, MD 20705, USA
[b] School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan 689-798, Republic of Korea
[c] School of Environmental Science and Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Republic of Korea

ARTICLE INFO

Article history:
Received 23 May 2015
Received in revised form 31 August 2015
Accepted 1 September 2015
Available online 8 September 2015
This manuscript was handled by Geoff Syme, Editor-in-Chief

Keywords:
Stressor–response
Machine learning
Temperature sensitivity
Chlorophyll-a

SUMMARY

To control algal blooms, the stressor–response relationships between water quality metrics, environmental variables, and algal growth need to be better understood and modeled. Machine-learning methods have been suggested as means to express the stressor–response relationships that are found when applying mechanistic water quality models. The objective of this work was to evaluate the efficiency of regression trees in the development of a stressor–response model for *chlorophyll-a* (Chl-a) concentrations, using the results from site-specific mechanistic water quality modeling. The 2-dimensional hydrodynamic and water quality model (CE-QUAL-W2) model was applied to simulate water quality using four-year observational data and additional scenarios of air temperature increases for the Yeongsan Reservoir in South Korea. Regression tree modeling was applied to the results of these simulations. Given the well-expressed seasonality in the simulated Chl-a dynamics, separate regression trees were developed for months from May to September. The regression trees provided a reasonably accurate representation of the stressor–response dependence generated by the CE-QUAL-W2 model. Different stressors were then selected as split variables for different months, and, in most cases, splits by the same stressor variable yielded the same correlation sign between the variable and the Chl-a concentration. Compared to physical variables, nutrient content appeared to better predict Chl-a responses. The highest Chl-a temperature sensitivities were found for May and June. Regression tree splits based on ammonium concentration resulted in a consistent trend of greater sensitivity in the groups of samples with higher ammonium concentrations. Regression tree models provided a transparent visual representation of the stressor–response relationships for Chl-a and its sensitivity. Overall, the representation of relationships using classification and regression tools can be considered a useful approach to assess the state of aquatic ecosystems and effectively determine significant stressor variables.

Published by Elsevier B.V.

1. Introduction

Excessive algal growth in freshwaters is globally recognized as a detrimental phenomenon. Excess algae can hamper navigation, deplete the oxygen stock in water, obliterate water clarity, cause the appearance of toxins, result in fish kills, promote growth of invasive algae species, ruin quality of surface waters for recreational use, and substantially decrease property values. As such, the monitoring and modeling of both algal growth and its physical and chemical controls constitutes an important part of environmental protection activities. Photosynthetic pigment content, including *chlorophyll-a* (Chl-a) concentrations, are measured as a surrogate for algal biomass because the cost and time required for Chl-a measurement is less than that for measurements of algal biomass. Chl-a is a response variable that is commonly used to measure biotic productivity that reflects the nutrient enrichment of a system. It is used in current numeric US EPA-approved criteria to indicate water impairment by contaminant levels of nitrogen and phosphorus (US EPA, 2013).

Trends and fluctuations of Chl-a concentrations can be used to reflect the corresponding trends and fluctuations of both the chemical parameters of water quality and physical environmental parameters. For example, inter-annual variations of Chl-a were about one order of magnitude in Lake Taihu in China (Xu et al., 2013), with oscillations greater than 1.5 orders of magnitude that were recorded during the 'June through October' periods in Oregon (Hoilman et al., 2008). Strong seasonality in Chl-a has also been observed in a multitude of lake monitoring studies, which is partially attributable to different limiting factors (Conley et al., 2009; Elser et al., 1990;

* Corresponding author. Tel.: +82 62 715 3277; fax: +82 62 715 2434.
E-mail address: joonkim@gist.ac.kr (J.H. Kim).

http://dx.doi.org/10.1016/j.jhydrol.2015.09.002
0022-1694/Published by Elsevier B.V.