Proceedings of the World Congress on New Technologies (NewTech 2015)
Barcelona, Spain – July 15 - 17, 2015
Paper No. 166

A Time Series Model for Prediction of Algae Concentration in River Systems

Sanghyun Kim
Department of Civil and Environmental Engineering/Pusan National University
Jangjun-dong san 30 Kumjungku, Busan, South Korea
kimsangh@pusan.ac.kr

Extended Abstract

Substantial blooming of algae has been one of primary water quality issues for heavily polluted inland water systems. Concentration of chlorophyll-a in river systems depends on various hydrometric and biochemical factors and corresponding mechanisms for growth and extinction are substantially complicate and prediction of algae concentration is challenging. A time series model can be an alternative tool for the modeling of chlorophyll-a concentration incorporating the stochastic structure in algae variation provided sufficient data exists. The variation of chlorophyll-a concentration in river systems shall be controlled through multiple processes such as hydro-meteorological, bio-chemical, and anthropogenic drivers. The temporal feature of different factors can be incorporated into distinctive pattern of chlorophyll-a variation. This means that the stochastic process to predict chlorophyll-a concentration is not necessary to be restricted into the traditional single process based approach such as autoregressive moving average model or seasonal autoregressive integrated moving average model. In order to improve predictability of existing model, an underlying assumption of existing modelling approaches is relaxed as the predicting variable is assumed to be determined through multiple independent processes. An enhanced stochastic model, namely a multiple process univariate model, is developed to address impact of distinct mechanisms associated with different drivers such as temperature, rainfall and nutrients.

Modelling of algae concentration time series follows four distinct procedures (e.g. statistical analysis and pretreatment, an identification of model structure, an estimation of parameters, and a diagnostic checking). If delineated model structure was found to be inappropriate in the checking procedure, modeling is restarted from the second step (model identification) to find better model structure. In order to estimate model parameters, the maximum likelihood estimates (Box and Jenkins, 1976) and conditional likelihood estimates were used to find the sum of the square surface of for a range of parameters and locate its minimum with the corresponding parameters. If multiple candidates were obtained through this procedure, a parsimony test was applied to select a final model. Akaike's information criterion (AIC) was used to find a balance between the variance of the residuals and the number of autoregressive and moving average parameters (Akaike, 1974).

Observations of algae concentration in 16 points for 4 major river systems in South Korea were used for modelling of chlorophyll-a concentration. Comparisons of modeling performance between traditional models and the proposed method show strengths of the multiple process univariate model both in predictability and parsimony in the model structure.

Akaike, H. (1974). A New Look At the Statistical Model Identification. *IEEE Transaction Automatic Control*, *19*(AC6), 716-723.

Box, G., Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control* (Revised Ed.). Englewood Cliffs, N. J.: Prentice-Hall.