5th World Conference on
Sustainability, Energy and Environment

Prague, Czech Republic 18 – 20 September 2025

SARIMAX Modeling of Chlorophyll-A In as Conchas Reservoir (Ourense, NW Spain)

María Pazo[1], Rafael da Hora[1], Mario Soilán[2], María Araújo[1]

[1]CINTECX, Universidade de Vigo, GESSMin Group, Vigo, 36310, Spain [2]CINTECX, Universidade de Vigo, Geotech Group, Vigo, 36310, Spain

ABSTRACT

The proliferation of cyanobacteria in water bodies is a critical environmental and public health concern, driven by nutrient enrichment and climate-induced changes. As toxic blooms become increasingly frequent, developing accurate predictive tools is essential for risk mitigation and environmental management. In this context, the present study proposes a Seasonal ARIMA with eXogenous regressors (SARIMAX) model to forecast chlorophyll-a concentrations in the eutrophic As Conchas reservoir (Ourense, NW Spain). Chlorophyll-a was selected as a key indicator of bloom intensity due to its strong correlation with cyanobacterial biomass. The model was trained using a four-year dataset (2017–2021) collected from three high-frequency sensors installed throughout the Limia River watershed. Environmental variables used as exogenous regressors in the SARIMAX model were selected based on their informational significance via a supervised Bayesian analysis, in which chlorophyll-a served as the target node. By incorporating exogenous regressors, the SARIMAX model significantly improved predictive performance compared to traditional ARIMA and SARIMA baselines. The optimal SARIMAX configuration achieved strong validation metrics, including $R^2 = 0.83$, MAE = 2.9%, RMSE = 6.0%, and an F1 score of 0.9. In conclusion, these findings underscore the potential of Bayesian and SARIMAX models as robust, data-driven tools for enhancing decision-making and supporting their integration into real-time water quality monitoring systems.

Keywords: Bayesian network; cyanobacteria; time-series modeling; SARIMAX; water quality.