PREDICTIVE MODELING AND SIMULATION SYSTEM FOR THE MANAGEMENT OF HARMFUL CYANOBACTERIA BLOOMS

Beatriz Herguedas-Pinedo
José L. Risco-Martín
Segundo Esteban
José A. López-Orozco
Eva Besada-Portas

Department of Computer Architecture and Automation
Universidad Complutense de Madrid
Calle Prof. José García Santesmases, 9
28040 Madrid, SPAIN
{behergue,jlrisco,sesteban,jalo,ebesada}@ucm.es

ABSTRACT

Water scarcity is increasing due to climate change, overexploitation and pollution. In addition, water bodies contain Harmful Cyanobacteria Blooms (HCBs) that produce toxins that are harmful to health, economy and environment. So far, these blooms have been assessed mainly by manual collection and analysis, or with the help of automatic instruments that acquire information from fixed locations. However, although having Early-Warning Systems (EWSs) to detect HCBs would be ideal, the procedures used do not usually provide data with sufficient resolution to anticipate their formation. Therefore, it is necessary to develop techniques and tools that combine data collection procedures with numerical simulations to detect, characterize, predict and respond to these outcrops. For this, it is proposed to implement a system for prediction and analysis of HCBs as part of an integral solution for its monitoring and management in real time, supported by a Model Based Systems Engineering (MBSE) infrastructure.

Keywords: Cyanobacteria monitoring, Harmful Cyanobacteria Blooms, Early Warning Systems, Data Analysis, Predictive Modeling.

1 INTRODUCTION

Water scarcity affects about 4 billion people (i.e., two-thirds of the world's population). This problem is closely linked to the term *water quality* (Cordy 2001), which refers to the chemical, physical and biological characteristics of water, which depends on its intended use. Nowadays, meeting quality standards is becoming increasingly difficult due to population growth, climate change, pollution and overexploitation, which introduce in the water bodies undesirable substances or microorganisms that can become a health risk.

One of those substances are cyanobacteria. They are unicellular microorganisms that use sunlight to photosynthesize, through which they manufacture their food from nutrients present in the water. Cyanobacteria are present in all environments and play an important role as primary producers. However, in warm and nitrogen-rich environments, they can multiply rapidly, creating blooms that spread throughout the water body and that are harmful to the health of humans, animals, or the environment by producing highly toxic secondary metabolites called cyanotoxins (**?**).

For this reason, a great deal of effort is invested in detecting these organisms. The most common procedure is to take samples manually for analysis to determine the presence of cyanotoxins that contaminate the water. This is time-consuming and economical ineffective due to the large amount of surface area that must be covered. The delay suffered by the results and the current limited capability to interpret these scenarios diminish the possibility of prediction, prevention, and mitigation of Harmful Cyanobacteria Blooms (HCBs) by the authorities (Meriluoto, Spoof, and Codd 2017).

To provide a comprehensive solution to the problem of monitoring and managing cyanobacteria, a novel model-driven Early-Warning System (EWS) is being implemented. It replicates the whole ecosystem of a modern EWS infrastructure, including monitoring agents (such as sensors or unmanned surface vehicles), control stations, predictive models, data analysis, training algorithms, etc. To better organize all the elements of the EWS, an Internet of Things (IoT)-based scheme is followed. To tackle such a system, the Discrete Event System Specification (DEVS) formalism (Zeigler, Muzy, and Kofman 2018) is used, as it provides a robust discrete event modeling and simulation framework to develop complex systems.

This paper focuses on presenting the prediction, training, and analysis models that are part of the mentioned infrastructure. As with the rest of the whole EWS, both the architecture design and its implementation are also based on the DEVS formalism. All this allows us to efficiently guarantee a system not only for detection but also for predicting and monitoring cyanobacteria outbreaks, in order to combate the risks that these microorganisms generate for health and the environment.

The remainder of this paper is organized as follows. Section 2 briefly explains the main topics of our research. Section 3 summarizes the EWS general architecture. Section 4 elaborates on the main elements of the implemented predictive support and analysis subsystem. Section 5 defines a case study and illustrates how the different elements work over it. Finally, Section 6 draws the conclusions of this work and propose some future lines of research.

2 BACKGROUND

This section introduces some basic notions used throughout this paper related to cyanobacteria and their impact, to early warning systems and their relevance, and to the DEVS formalism.

2.1 Cyanobacteria

Cyanobacteria live in water environments, obtaining energy through photosynthesis. Their growth is favored in calm waters and high temperatures. These organisms vary in morphology from unicellular and filamentous to colonial forms (Aguilera, Klemenčič, Sueldo, Rzymski, Giannuzzi, and Martin 2021). This group of bacteria often dominates the upper layers of microbial mats found in many water environments. However, due to their strong light requirements, they often migrate vertically within the mat to find the location that provides the right level of light for photosynthesis. Finally, they are also displaced horizontally due the water currents and winds.

Some cyanobacteria form colonies by clustering into microbial mats that perform important ecological functions. However, they sometimes generate toxic secondary metabolites that are harmful for the environment, deplete oxygen from the water, or reduce sunlight penetration, compromising therefore the development of other species (Mullineaux and Wilde 2021). The most common methods currently used to deal with cyanobacteria are chemical treatments and the reduction of nutrients entering water systems. However, they are neither easy to carry out nor very effective.

2.2 Early-Warning Systems

It is essential to implement EWSs that collect information on the state of the water, identify the presence of inadequate levels of cyanobacteria, and facilitate the prediction of HCBs formation and their management. EWSs are currently being developed as predictive systems (Wu, Zhang, Hou, Tian, Chu, and Wang 2022), supported by a set of *base models* that describe the behavior of HCBs. These models are integrated into more generic software tools such as EEMS (DSI 2022), or MIKE Powered by DHI (2022).

Data-driven base models, such as Artificial Neural Networks (Wei, Sugiura, and Maekawa 2001), Bayesian Networks (Moe, Couture, Haande, Lyche Solheim, and Jackson-Blake 2019) and Genetic Programming (Sivapragasam, Muttil, Muthukumar, and Arun 2010), can be found in the literature. Mathematical models, such as Eulerian Models (Vinçon-Leite and Casenave 2019), Lagrangian Models (Van Sebille, Griffies, Abernathey, Adams, Berloff, Biastoch, Blanke, Chassignet, Cheng, Cotter, et al. 2018), Rule-Based Models (Kim, Cao, Jeong, Recknagel, and Joo 2007) or Probabilistic Models (Zhao, Shao, Yang, Ren, Ge, Feng, Dong, and Zhao 2019), are also used. Finally, there are some Mechanistic Models based on ecosystem knowledge (Rousso, Bertone, Stewart, and Hamilton 2020), which use the correlation between phycocyanin fluorescence and cyanobacteria biomass in water (Izydorczyk, Tarczynska, Jurczak, Mrowczynski, and Zalewski 2005). All of them can be used to obtain information on the modeled phenomena and can be used to identify probable points of HCBs growth.

2.3 DEVS

DEVS is a formalism for modeling discrete event systems based on set theory (Zeigler, Muzy, and Kofman 2018). It includes two model types, atomic and coupled, with an interface consisting of input and output ports to communicate with other models.

Atomic models have a state with a parameter σ, which determines the duration the state remains unchanged. Once the time assigned to the state has passed, an internal transition is triggered, producing a local state change. At that time, an event can be generated and distributed through the output ports of the model. External input events (received from other models) are collected in the input ports. An external transition function specifies how to react to those inputs.

Coupled models are the aggregation/composition of two or more models (atomic or coupled), connected by explicit couplings. This makes DEVS closed under coupling and allows us to use networks of systems as components in larger coupled models, leading to hierarchical and modular constructions.

Finally, once a system is described according to DEVS theory, it can be easily implemented using one of the many DEVS Modeling and Simulation (M&S) engines that have come into existence in the last decades.

3 GENERAL ARCHITECTURE OF THE EARLY WARNING SYSTEM

The modeling framework created to simulate and operate our EWS, which is depicted in Fig. 1 and aimed at real-time monitoring and prediction of cyanobacteria blooms, has been designed following the principles of Model Based Systems Engineering (MBSE). It is developed using the well-known DEVS modeling and simulation formalism, in order to properly organize its whole structure and coarse-grain behavior. At the same time, the fine-grain prediction, training and data analysis (as well as the properties of other parts of

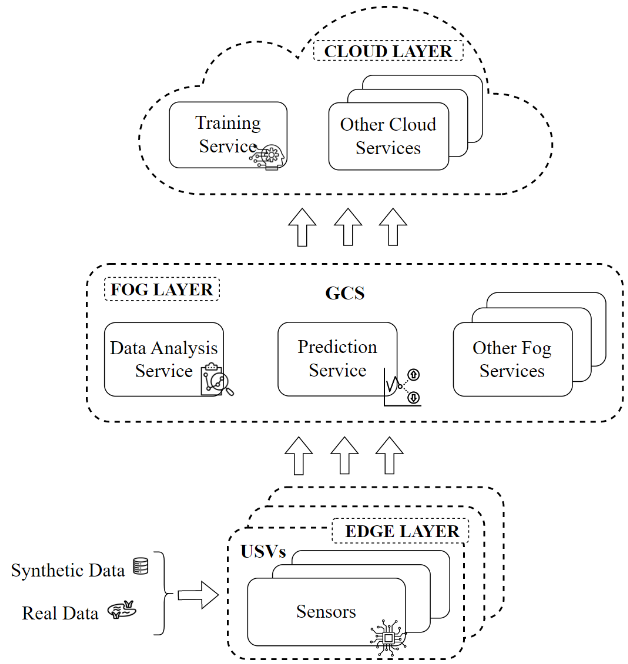

Figure 1: EWS general architecture.

the system) are handled by the transition functions and encapsulated into the aforementioned *base models* or other general algorithms related to data analysis.

The EWS model is divided into the three classic IoT layers: Edge, Fog, and Cloud. Water bodies are monitored in the Edge layer by a set of sensors, included those ones on board Unmanned Surface Vehicles (USVs), which continuously send data to the Ground Control Station (GCS). In the Fog layer, experts can analyze the data, run HCB predictions or plan the trajectories of the USVs using some of the implemented base models and algorithms. In addition, the framework supports horizontal scalability, allowing the addition of more water bodies and more USVs with the support of the Cloud layer, where authorities can study reports and make high-level decisions supported by the EWS. In addition, the EWS is fed with data that can come from an actual water body or a synthetic database.

The Edge layer contains the USV model, which encapsulates the sensors' atomic models. These sensors measure the different parameters that affect the growth of cyanobacteria. Specifically, the USV model has currently a positioning unit that calculates the coordinates of the USV and nine sensors created as Digital Twins (DTs), which can take data from actual sensors or be replaced by their physical counterparts. The Fog layer contains a series of atomic models. The main one is the GCS, which represents the core of the computer infrastructure of the ground control station. As mentioned above, this layer contains several atomic models in charge of running different services through their transition functions, such as prediction or data analysis. Finally, the Cloud model can also run additional services, including training the predictive models or generating reports.

4 PREDICTION, TRAINING AND ANALYSIS SYSTEMS

In the following, we describe and elaborate on those models of the EWS architecture that are directly related to the predictive and training HCB of the cyanobacteria blooms, and to the generation of data reports.

4.1 Bloom Prediction

The primary function of the prediction subsystem is to calculate the chances of formation and location of blooms in each water body. For this purpose, the following system of differential equations is used. This base model represents the HCB density and position as a function of certain factors, such as the amount of dissolved nitrogen and oxygen or the water velocity:

$$\frac{dr(t)}{dt} = k_{growth} \cdot u(t) - k_{decline} \cdot (r(t) - r(0)) \tag{1}$$

$$\frac{dx(t)}{dt} = k_{speed} \cdot v_x(t) \tag{2}$$

$$\frac{dy(t)}{dt} = k_{speed} \cdot v_y(t) \tag{3}$$

In particular, in Eq. (1), r represents the bloom density, while u is the concentration of nutrients present in the water. Besides, in Eq. (2) and (3), x and y are the coordinates of the position of the bloom at a given height, whereas v_x and v_y are the water velocity at the same coordinates. Regarding the constants in those equations, k_{growth} is the HCB growth constant, $k_{decline}$ is the decay constant, and k_{speed} represents the percentage of the water velocity that is transferred to the HCB.

The nutrient concentration u is calculated as follows:

$$u(t) = k_{photo} \cdot photo(t) + k_{breath} \cdot breath(t), \tag{4}$$

where $breath$ represents respiration and $photo$ represents photosynthesis, and k_{breath} and k_{photo} are the constants that measure the percentage of respiration and photosynthesis that effectively affect the amount of nutrients dissolved in the water.

The value of all the constants of the model are obtained by training the system, which will start with some given values for them. As shown in the following section, these values are adjusted each time the error between actual value and simulated value exceeds a given threshold.

The previous base model is encapsulated into a DEVS atomic model. The purpose is to verify and validate the whole EWS and then improve the base models keeping the same inputs and outputs in the DEVS components. In this case, the internal transition includes the model of differential equations, which is solved using the *odeint* function from the Livermore Solver for Ordinary Differential Equations (LSODE) library. This library uses a predictor-corrector Adams method that combines an explicit variant with an implicit one. The explicit variant predicts an approximation, while the implicit variant corrects it. The DEVs atomic model that encapsulates the previous equations receives the data transmitted by the sensors related to the variables u, v_x and v_y. Once the bloom density is calculated with Eq. (1), if the density is sufficient to consider that there is massive upwelling, its new position is calculated with Eq. (2) and (3). If not, the position is reset to the initial search position.

4.2 Training System

The training service aims to retrain the bloom predictive model to improve the results of the new predictions. Specifically, we want to adjust the parameters of Eqs. (1) and (4) to incorporate the behavior of the HCB under analysis in case the predictive model did not work well. This is achieved by minimizing the error, i.e., the distance between our prediction and the actual bloom density data (measured by the USV) at the position calculated by Eq. (2) and (3).

The combination of Eq. (1) and Eq. (4) model the bloom density evolution as a function of 4 constants: k_{growth}, $k_{decline}$, k_{breath} and k_{photo}, that can be translated by three parameters providing the system with three degrees of freedom, as can be seen in the following expression obtained substituting Eq. (4) in (1), and where $k_1 = k_{growth} \cdot k_{photo}$, $k_2 = k_{growth} \cdot k_{breath}$ and $k_3 = k_{decline}$:

$$\frac{dr(t)}{dt} = k_1 \cdot photo(t) + k_2 \cdot breath(t) - k_3 \cdot (r(t) - r(0)) \tag{5}$$

The training model must adjust these constants to minimize the error in the bloom density when this error increases substantially. For this purpose, we use the function *curve_fit* from the *scipy* library, which uses Nonlinear Least Squares (NLS) to fit the predictive model to the data. That is, given a numerical approximation to the differential equation evaluated at time instant t_i, and an initial $\mathbf{k}_0$ value of $\mathbf{k} = [k_1, k_2, k_3]$, the set of parameters to fit are estimated by minimizing the following expression:

$$\hat{\mathbf{k}} = \min \sum_{i=1}^{N} [r_i - \hat{r}(t_i, \mathbf{k}, r_0)]^2 \tag{6}$$

where $\hat{\mathbf{k}}$ represents the parameters to be estimated, r_i is the actual bloom density data at instant t_i, and $\hat{r}(t_i, \mathbf{k}, r_0)$ represents the prediction for instant t_i (Carey and Ramsay 2021).

The training system is not encapsulated in a DEVS atomic model. Instead, the EWS coupled model is simulated replicating the current mission, and if the error exceeds a given threshold, the training service is launched as follows:

```
# Initializes the DEVS simulation with the EWS model
Coodinator coordinator(ews_model)

# Run the simulation and get the error:
coordinator.simulate()
error = ews_model.get_predictive_error()

# Adjust parameters when error > threshold
if (error > threshold):
   ews_model.adjust_predictive_model()
```

4.3 Data Analysis

The analysis subsystem, distributed among the Fog and Cloud layers, has the task of analyzing the information stored in the database and generating reports.

Fog layer reports consist of different graphs that help to visualize everything that happened during the simulation. They consist of 22 graphs, 6 of them dynamic. Each report in this layer is divided into five main sections. The initial one includes dynamic charts. Here the domain expert can get a visual idea of the simulation, showing maps with the actual measurements of oxygen, nitrogen, water and wind velocity, and cyanobacteria density, as well as the USV trajectory and a perimeter marking the location of the prediction made by the model. The second section shows the evolution of sensors' measurements: water temperature and speed, oxygen and nitrogen density, etc. The third section shows the same information but with the predicted bloom variables, indicating the bloom density and whether a massive upwelling is detected. The fourth section shows the evolution of the USV battery status and velocity. Finally, the fifth section offers the solution of the predictive model that allows us to observe the error in the prediction.

The Cloud layer reports consist of a table showing the summary of the simulation results, with the arithmetic mean and standard deviation of the most important parameters of the EWS like battery status, solar radiation, water temperature, etc. On the other hand, it depicts a heat map showing the locations where blooms have been often found. These locations are referred to as *hot spots* because there is a significantly high probability that HCBs will reappear there.

5 RESULTS

This section presents a simulation scenario that tests the architecture and implementation of the developed models, showing simulation results.

5.1 Scenario

The water body selected for the simulation is the southern area of Lake Washington, located in the state of Washington (USA). In this case, the water body dataset is completely synthetic and is generated with EEMS, which incorporates an accurate lake model and allows the simulation of cyanobacteria. The procedure will be similar when we have access to actual data obtained by the USV that we are building.

In particular, the simulation takes place with the data available from EEMS, between August 23, 2008 at $00:00$ hours, and September 7, 2008 at the same time. The prediction model initializes the bloom location at the beginning of the day in the southern part of the lake, next to a channel that flows into a shallow area that provides ideal conditions for the formation of upwelling and nitrate accumulations. During the simulation period, every 30 minutes, the sensors take measurements and send the data to the GCS. In turn, the GCS sends the information to the prediction service, which uses the data received from the sensors related to water velocity, oxygen and nitrate concentrations, and solar radiation to predict the appearance and displacement of the HCB. Afterwards, it sends the prediction back to the GCS, which calculates and sends to the USV the next point on its trajectory. To close the loop, the USV navigates and takes measurements at that new position (in the simulation, it does it virtually), sending the information back to the Fog layer and launching the repetition of the whole process described above. Finally, before ending the simulation, the GCS sends all the information to the Cloud layer and activates the analysis DEVS model. This model generates the reports, usable to evaluate the possibility of readjusting the parameters and improving the accuracy of the predictive model.

5.2 Fog Analysis Reports

The Fog layer report shows all the simulation details, generating a video ((Herguedas 2023)), which includes 6 dynamic graphics that illustrate the simulation comprehensively and visually.

Figure 2 shows an image taken from a frame of the simulation video. On one hand, the upper left graph shows the map of cyanobacteria density in the surface layer, limits the position and size of the predicted upwelling with a red circle on the map, and indicates with a black dot the current location of the USV. The map representing the cyanobacteria density represents data from the EEMS simulator to validate that the predictive model correctly anticipates the dynamics of the HCB. Besides, the top center plot shows the complete map of dissolved oxygen density in the surface layer, while the upper right graph shows a similar map with the nitrate density. On the other hand, the lower left graph shows three signals as a function of the time of day: (i) the solar radiation measured by the irradiance sensor is plotted in blue, (ii) the water temperature measured by one of the USV sensors is plotted in red; and (iii) USV battery level is plotted in black. The remaining two lower plots show the map with velocities of the water (center) and of the wind (right) in the water surface.

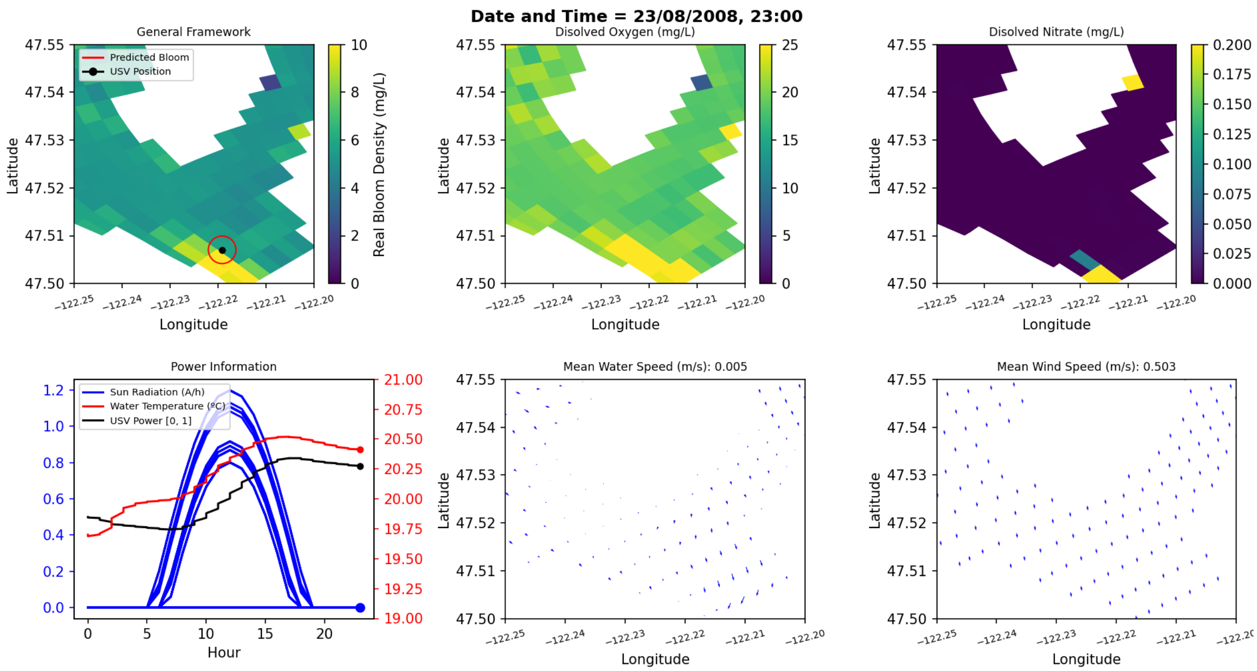

Figure 2: General simulation framework.

In addition, the Fog report generates a series of graphs representing the evolution of different data used by the predictive model to forecast bloom formation throughout the simulation. Figure 3 contains 8 of these plots, showing different measurements over time: such as water temperature and velocity, the amount of nitrate and dissolved oxygen in the water, the predicted bloom density, the detection of massive upwelling, and the USV battery and its velocity.

Finally, the first graph of Figure 4 shows the values of the actual bloom density and its prediction, while the second graph in that Figure shows the difference between both values through the solid line and its dashed line represents the arithmetic mean of the error. This information is necessary to decide if the predictive model must be retrained (as we will see later in Section 5.4).

5.3 Cloud Analysis Report

On one hand, the Cloud layer reports summarize in a table the most relevant characteristics of different simulation variables, such as USV battery, solar radiation, water temperature, and bloom density. For this purpose, various operations are performed on the data, such as arithmetic mean, standard deviation, fiftieth percentile, minimum and maximum. This information can be seen in Table 1.

On the other hand, an interactive heat map is also generated at the Cloud layer, showing the most likely areas to find HCBs based on the measurements made during the simulation. This map, which shows in lighter colors the hottest spots, can be seen in Fig. 5. As the map is zoomed in, the distribution of the most probable areas where cyanobacterial blooms may appear can be seen more clearly.

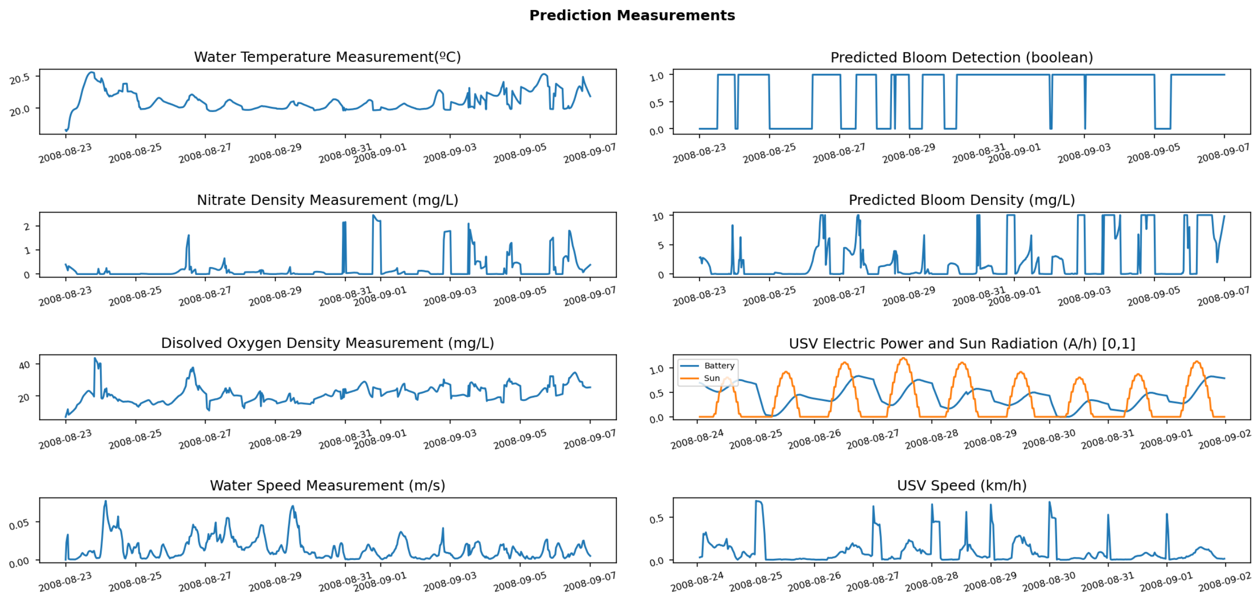

Figure 3: Prediction Measurements Graphic over the simulation time.

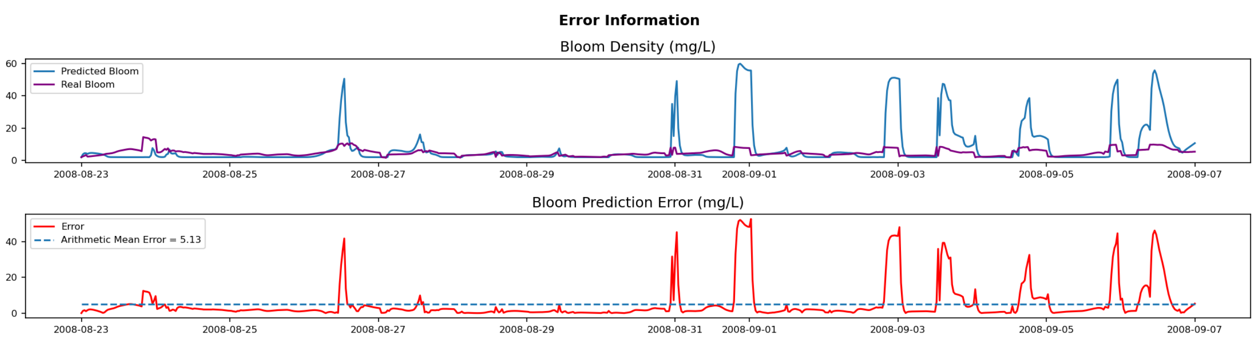

Figure 4: Predicted Bloom Density Error.

5.4 Retraining the model

Once the simulation is finished and the data analysis model is executed, we can see the prediction error in the final section of the Fog layer report. This first simulation was carried out with the default values of the constants, which are $k_1 = 1$, $k_2 = 1$ and $k_3 = 1$.

The result of such simulation and the difference between the actual and predicted data for bloom density was already displayed in Fig. 4. The arithmetic mean error for this simulation turned out to be 5.13, which is high. Thus, the training service is triggered, which starts the nonlinear least squares method to minimize the error. When the training service finishes, it returns the new values of the 3 params of the density base model, which in this case are $k_1 = 1.02424907$, $k_2 = 0.35783188$ and $k_3 = 2.2402504$, and which are substituted by the old ones in the differential equations. Thus, a new simulation run is relaunched with a more accurate predictive model, as can be seen in Fig. 6, which depicts the new difference between predicted.

Figure 7 depicts the before and after training predictions, compared against the actual bloom density data. The density line of the adjusted models follow better the data dynamics than the line of the unadjusted prediction. Besides, the arithmetic mean error between the prediction and the real data has decreased from 5.13 to 2.09 with respect to the previous run, implying that the error has been reduced by more than half.

	USV Electric Battery (A/h)	Solar Radiation (A/h)	Water Temperature (ºC)	Bloom Density (mg/L)
Arithmetic Mean	0.919	0.013	20.169	0.752
Standard Deviation	0.102	0.015	0.370	1.720
50th Percentile	0.952	0.000	20.262	0.327
Minimum	0.445	0.000	19.357	0.009
Maximum	1.000	0.040	20.585	10.000

Table 1: Table for Data Analysis.

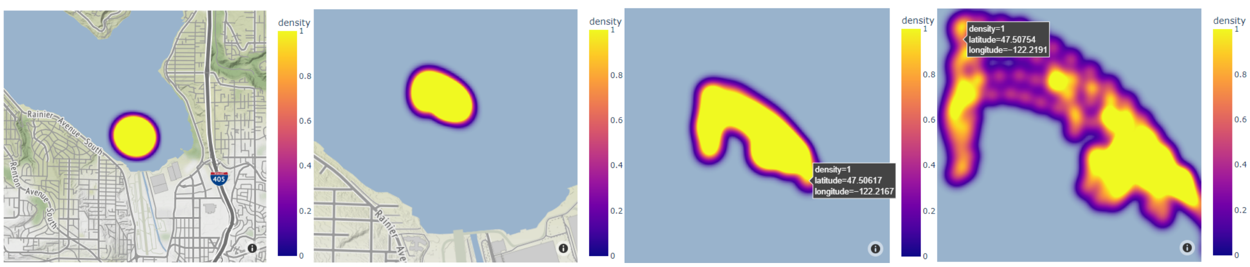

Figure 5: Heat map of the areas most likely to find bloom.

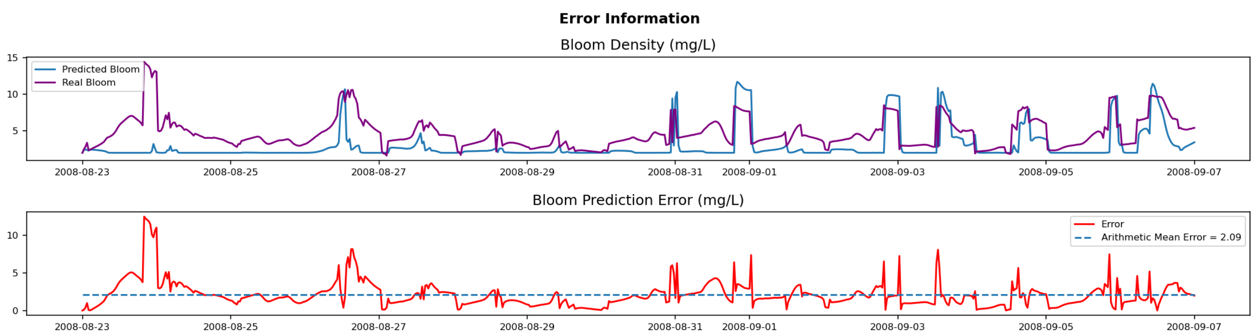

Figure 6: Predicted Bloom Density Error after training.

However, it is worthwhile to mention that although the density predictive model does not yet follow the same dynamics as the real data, it is already able to detect the peaks of the dynamics.

6 CONCLUSIONS AND FUTURE WORK

A model-driven methodology can help domain experts and authorities design comprehensive early warning systems. The proposed EWS framework facilitates the development of a sustainable and efficient bloom management system by allowing the integration of different models that collaborate for the same purpose. Within this framework, the predictive model enables to anticipate HCBs, the reports asses the situation of the water body, and the training subsystem improve the perception models.

Some future work in the prediction subsystem, is already planned. First, we want to incorporate more sophisticated behaviors in the bloom behavior (e.g. applying fluid dynamics to nutrient transport, or incorporating the HCB vertical movement into the model) and test other numeric integration methods. Besides, the base model could detect hotspots instead of replicating the HCB dynamic, by using Artificial Intelligence-based models. These new types of models will be straightforward to add into the whole system, thanks to DEVS characteristics. Regarding the data analysis models, we are contacting domain experts and authorities to

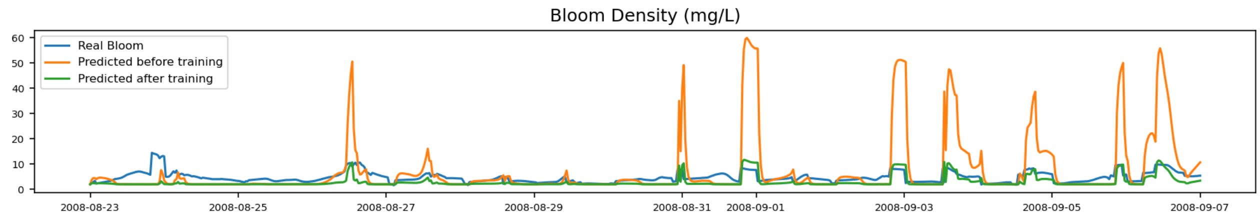

Figure 7: Differences in bloom prediction before and after training

detail in the tables and the graphs their special necessities. Finally, we need real data to test and train the models, since we have not yet finished testing the behavior of our real USV and sensors. However, it is worth noting that the use of synthetic data is allowing us to design our EWS and will help authorities in the future to gain confidence in the system without really acting over it.

ACKNOWLEDGMENTS

This work has been supported by the Research Project IA-GES-BLOOM-CM (Y2020/TCS-6420) of the Synergic program of the Comunidad Autónoma de Madrid, and by the Research Project SMART-BLOOMS (TED2021-130123B-I00), funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR.

REFERENCES

Aguilera, A., M. Klemenčič, D. J. Sueldo, P. Rzymski, L. Giannuzzi, and M. V. Martin. 2021, 03. "Cell Death in Cyanobacteria: Current Understanding and Recommendations for a Consensus on Its Nomenclature". *Frontiers in Microbiology* vol. 12.

Carey, M., and J. Ramsay. 2021. "Fast stable parameter estimation for linear dynamical systems". *Computational Statistics & Data Analysis* vol. 156, pp. 107124.

Cordy, G. E. 2001. *A Primer on Water Quality*. US Geological Survey.

DHI 2022. "MIKE ECO Lab". (Visited on 11/2022).

DSI 2022, January. "EE Modeling System". https://www.eemodelingsystem.com.

Beatriz Herguedas 2023. "DEVS BLOOM Simulation".

Izydorczyk, K., M. Tarczynska, T. Jurczak, J. Mrowczynski, and M. Zalewski. 2005. "Measurement of phycocyanin fluorescenceas an online early warning system for cyanobacteria in reservoir intake water". *Environmental Toxicology* vol. 20 (4), pp. 425–430.

Kim, D.-K., H. Cao, K.-S. Jeong, F. Recknagel, and G.-J. Joo. 2007. "Predictive function and rules for population dynamics of Microcystis aeruginosa in the regulated Nakdong River (South Korea), discovered by evolutionary algorithms". *Ecological Modelling* vol. 203 (1), pp. 147–156. Special Issue on Ecological Informatics: Biologically-Inspired Machine Learning.

Meriluoto, J., L. Spoof, and G. A. Codd. 2017. *Handbook of cyanobacterial monitoring and cyanotoxin analysis*. John Wiley & Sons.

Moe, S. J., R.-M. Couture, S. Haande, A. Lyche Solheim, and L. Jackson-Blake. 2019. "Predicting Lake Quality for the Next Generation: Impacts of Catchment Management and Climatic Factors in a Probabilistic Model Framework". *Water* vol. 11 (9).

Mullineaux, C. W., and A. Wilde. 2021, 06. "Bacterial Blooms: The social life of cyanobacteria". *eLife* vol. 10.

Rousso, B. Z., E. Bertone, R. Stewart, and D. P. Hamilton. 2020. "A systematic literature review of forecasting and predictive models for cyanobacteria blooms in freshwater lakes". *Water Research* vol. 182, pp. 115959.

Sivapragasam, C., N. Muttil, S. Muthukumar, and V. Arun. 2010. "Prediction of algal blooms using genetic programming". *Marine Pollution Bulletin* vol. 60 (10), pp. 1849–1855.

Van Sebille, E., S. M. Griffies, R. Abernathey, T. P. Adams, P. Berloff, A. Biastoch, B. Blanke, E. P. Chassignet, Y. Cheng, C. J. Cotter et al. 2018. "Lagrangian ocean analysis: Fundamentals and practices". *Ocean Modelling* vol. 121, pp. 49–75.

Vinçon-Leite, B., and C. Casenave. 2019. "Modelling eutrophication in lake ecosystems: a review". *Science of the Total Environment* vol. 651, pp. 2985–3001.

Wei, B., N. Sugiura, and T. Maekawa. 2001. "Use of artificial neural network in the prediction of algal blooms". *Water Research* vol. 35 (8), pp. 2022–2028.

Wu, Y., J. Zhang, Z. Hou, Z. Tian, Z. Chu, and S. Wang. 2022. "Seasonal Dynamics of Algal Net Primary Production in Response to Phosphorus Input in a Mesotrophic Subtropical Plateau Lake, Southwestern China". *Water* vol. 14 (5).

Zeigler, B. P., A. Muzy, and E. Kofman. 2018. *Theory of modeling and simulation: discrete event & iterative system computational foundations*. Academic press.

Zhao, C., N. Shao, S. Yang, H. Ren, Y. Ge, P. Feng, B. Dong, and Y. Zhao. 2019. "Predicting cyanobacteria bloom occurrence in lakes and reservoirs before blooms occur". *Science of The Total Environment* vol. 670, pp. 837–848.

AUTHOR BIOGRAPHIES

BEATRIZ HERGUEDAS-PINEDO is a researcher at the Complutense University of Madrid (UCM), Spain. She holds a BS degree in Mathematics and Computer Science Engineering and a MS degree in Computer Science Engineering, all from the Complutense University of Madrid. Her research interests include systems modeling and simulation. Her email address is behergue@ucm.es.

JOSÉ L. RISCO-MARTÍN received his Ph.D. from UCM, where he currently is Full Professor in the Department of Computer Architecture and Automation. His research interests include systems modeling, simulation, and optimization. He can be reached at jlrisco@ucm.es.

SEGUNDO ESTEBAN is an Associate Professor of Systems Engineering and Automation at UCM. He holds a Ph.D. in Physics from the same University. His research interests include Systems Modeling and Control. His email address is sesteban@ucm.es.

JOSÉ A. LÓPEZ-OROZCO is a Full Professor in the UCM. He holds a Ph.D. in Physics from the same University. His research interests include multisensor data fusion, control and planning of unmanned vehicles, and robotics. His email address is jalo@ucm.es.

EVA BESADA-PORTAS is an Associate Professor of Systems Engineering and Automation at UCM. She also holds a PhD in Computer Systems from UCM. Her research interests include uncertainty modeling and simulation, optimal control and planning of unmanned vehicles. Her email address is ebesada@ucm.es.