VIRTUAL SENSORS TO MINIMIZE RISK OF WATER REUSE IN FOOD INDUSTRY

Batlle Ribas M.[1], Cros Herrero J.[1], and Queraltó López A.[1]

[1]*Adasa Sistemas S.A.U,*

ABSTRACT

Water scarcity, exacerbated by the ongoing climate crisis, necessitates efficient water reuse strategies, particularly in water-intensive industries such as food and beverage. This abstract discusses the implications of drought in Catalonia on water consumption in the meat sector, current legislation governing water reuse, and the application of advanced monitoring technologies to ensure water quality and safety.

1. Introduction

Water scarcity, intensified by climate change, makes water reuse essential for industries like food and beverage, especially the water-intensive meat sector. The prolonged drought in Catalonia caused the declaration of the Drought Emergency State on February1, 2024, which entailed a 25% reduction in industrial water usage, underscores the urgency of this issue. The REAQUA (BETA s.f.) project identified that at least 15% of water in slaughterhouses, mainly for cleaning trucks and corrals, can be reused. To ensure water quality, aquaBio technology monitors *E. coli* and coliforms, meeting regulatory standards.

In the VECOLI project, a virtual sensor was developed to complement aquaBio, offering real-time *E. coli* predictions and alerting users to water quality issues. Validated in a slaughterhouse, this system integrates artificial intelligence for low-cost, continuous monitoring. It ensures safe reclaimed water use, adheres to regulations, and can be scaled for broader applications.

Spanish RD1620/2007 regulates the reuse of treated water, using *Escherichia coli* as a key indicator of microbiological quality. However, the food industry faces stricter regulations, requiring special authorization from health authorities. This makes risk assessment and obtaining continuous water quality approvals more challenging, especially in demonstrating the consistent suitability of reused water for industrial processes and cleaning.

The VECOLI virtual sensor offers real-time *E. coli* monitoring, alerting users to water quality changes and ensuring safe water reuse for cleaning. Powered by AI, this low-cost, automated system processes physical and chemical data, continuously retraining with *E. coli* measurements from aquaBio, preventing the use of compromised water and protecting public health.

2. Methodology

The functional requirements for the solution involve creating a virtual sensor capable of estimating *E. coli* values in real time, based on aquaTest measurement equipment, which includes sensors for various parameters such as temperature, pH, Oxidation-Reduction Potential (ORP), conductivity, dissolved oxygen, turbidity, UV254 absorbance and 365nm absorbance, as well as a probe for microbial activity BES. It also incorporates aquaBio equipment specifically for measuring *E. coli*, though these measurements are discrete rather than real-time, and are employed for training and verification of the system. Data is transmitted via a communication device to an acquisition and visualization system, responsible for generating the virtual *E. coli* sensor, as well as analyzing and representing the information. It also supports data acquisition from various external sources, such as meteorological data, and manual input of laboratory data. The system allows the creation of configurations for data preprocessing, model training, and result inference, with the capability to perform these operations on user demand.

It also enables the comparison of different artificial intelligence models and includes a frontend for real-time data visualization, display of model weights, and generates quality assessment reports for the virtual sensor's results.

The elements of the system installed are represented in Figure 1:

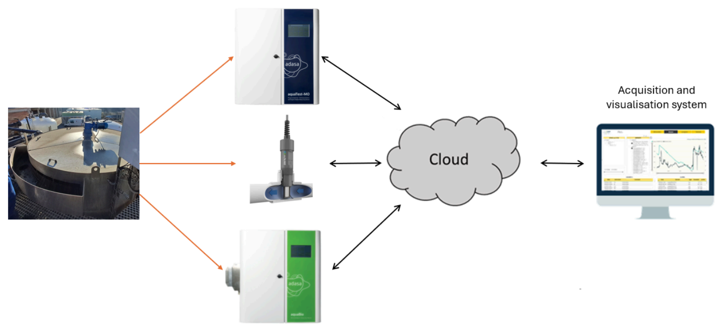

Figure 1 Schematic of aquaBio and VECOLI technology integration

The virtual sensor system integrates several components into a cohesive structure, ensuring efficient data handling and accurate predictions. It includes a web interface for access and employs Docker for containerization to maintain modularity. The system features three primary modules:

1. **Data Preprocessing**: This module manages missing values, filters outliers with a moving average technique, and generates additional variables to capture data patterns. Preprocessed data is stored in a database.
2. **Model Training**: Utilizing preprocessed data, this module trains predictive models using algorithms like linear and polynomial regression, neural network (MLP regressor) and decision tree-based models (XGBoost). Models are evaluated with training and validation datasets, and the best-performing model is stored for future predictions.
3. **Model Predictions and Validation**: Newly received data are processed similarly to the training phase, and the selected model generates predictions which are validated against measured data.

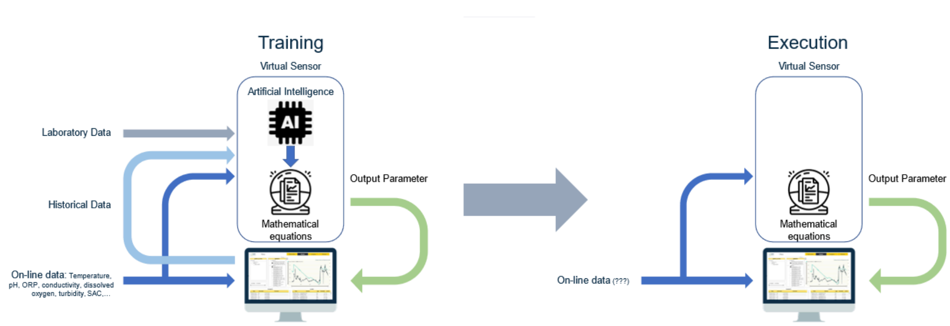

Figure 2 Schematic of aquaBio and VECOLI technology integration

3. Result & Discussion

The system was installed at the outlet of the wastewater treatment plant at Avinyó’s slaughterhouse. The sensor was trained with different mathematical models: Second degree polynomial, lineal regression, XGBoost (combined decision trees) and MLP regressor (neural network) for two different periods of time: The first corresponding to the outlet of the textile filter (from March to June) of the Wastewater treatment plant a Avinyó and a second corresponding to the same sample after the addition of chlorination (June and July). In addition, and to evaluate the result of the model, the real value of *E. coli* provided by the aquaBio measuring device and the virtual sensor were compared. Table 1 shows the results of accuracy in the prediction and error in the measure when aquaBio and virtual sensor are compared for the different periods of training and measure of the data: a) training with data from 03/2024 to 05/2024 and error in measuring from 03/2024 to 05/2024; b) training from 03/2024 to 05/2024 and measuring from 03/2024 to 07/2024; and c) training from 03/2024 to 07/2024 and measuring from 03/2024 to 07/2024. The results of accuracy in the prediction and error in the measure when aquaBio and virtual sensor are compared.

Table 1 Comparison of the results of the different models of the estimated values of the considering different periods of training and measure.

	Training from 03/2024 to 05/2024 and measure from 03/2024 to 05/2024		*Training from 03/2024 to 05/2024 and measure from 03/2024 to 07/2024*		*Training from 03/2024 to 07/2024 and measure from 03/2024 to 07/2024*	
Model	**Error**	**Accuracy**	**Error**	**Accuracy**	**Error**	**Accuracy**
Polynomial	3.12%	77.51%	17.20%	77.51%	7.87%	76.2%
XGBoost	4.26%	86.73%	13.18%	86.73%	7.97%	82.33%
Linear regression	6.15%	89.92%	12.53%	89.92%	11.05%	84.44%
MLP Regressor	4.92%	92.25%	11.73%	92.25%	7.9%	77.49%

Despite the large training accuracy of all models, with values between 76.2 and 92.25%, those with lower overall prediction error are the polynomial and the MLP regressor. If we focus in the polynomial model, the graphical results for these three periods, are shown in Figure 3: Training with data from 03/2024 to 05/2024 and error in measure from 03/2024 to 05/2024 (Figures 3 left), training from 03/2024 to 05/2024 and measure from 03/2024 to 07/2024 (Figure 3 middle), and training from 03/2024 to 07/2024 and measure from 03/2024 to 07/2024 (Figure 3 right).

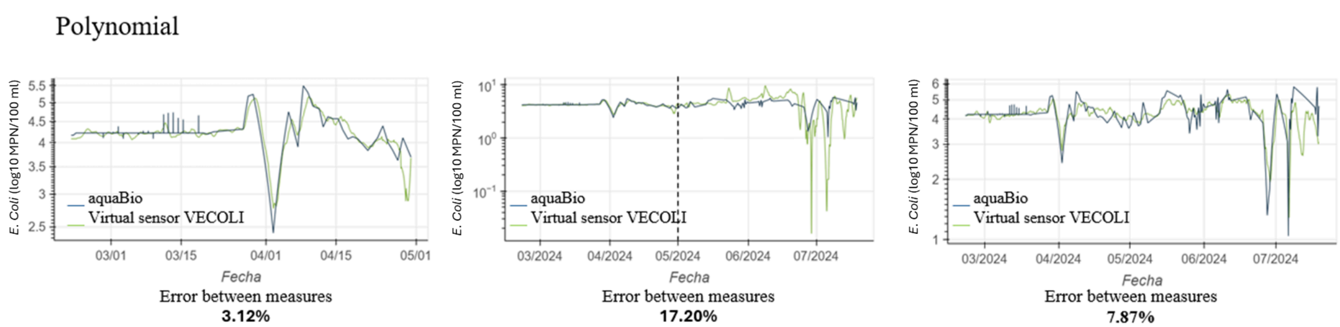

Figure 3 Graphical representation of the results for polynomial model in the three different periods of training and measure. The dashed line shows the results of the prediction without the use of the data for the training.

In a predictive model, certain parameters carry more weight than others, and these are ranked based on their relative importance. The ranking reflects their contribution to the model's accuracy in prediction. In Figure 4, the results for the polynomial model are presented for the two different training periods. These figures highlight how certain parameters have a greater impact on the model's predictive accuracy, with the parameters ranked according to their importance in the prediction process.

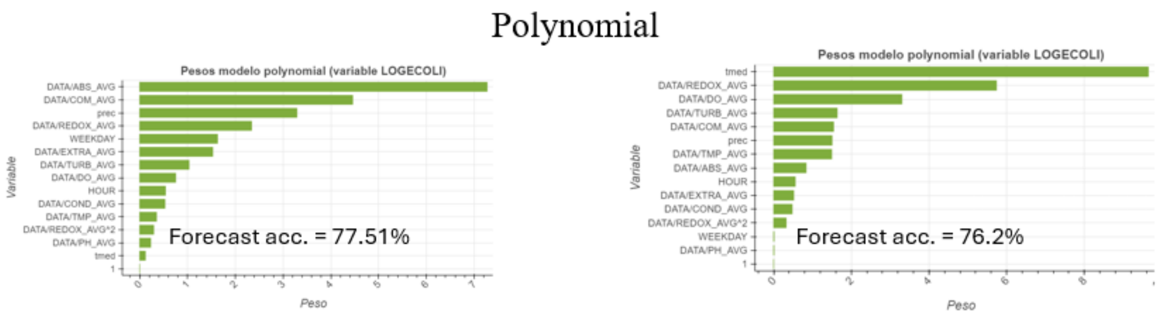

Figure 4 Ranking of the parameters contributing to the polynomial prediction system for the two-training periods. Left: from 03/2024 to 05/2024; Right from 03/2024 to 07/2024.

In the case of *E. coli*, the training results for the polynomial model with all data (Figure 4 right) show that the most influential variables are temperature (tmed), REDOX (REDOX_AVG), dissolved oxygen (DO_AVG). Variables with moderate influence include turbidity (TURB_AVG), compensation at 365 nm (COM_AVG). rain (prec), temperature of the sample (TMP_AVG) and absorbance at 254 nm (ABS_AVG). While variables such as time (HOUR), BES (EXTRA_AVG), and conductivity (COND_AVG) have low influence, and the day of the week (WEEKDAY) and pH (PH_AVG) have nearly null effect.

4. Impact and Scalability

The successful implementation of VECOLI showcases the potential to scale and replicate this system in other regions and sectors. This innovation not only enhances water reuse practices but also contributes to the resilience of the meat industry and overall water resource management.

5. Conclusion

Addressing water scarcity through advanced water reuse technologies is vital for industrial sustainability, especially in drought-prone regions like Catalonia. The integration of systems like aquaBio and VECOLI provides a robust framework for ensuring water quality, meeting regulatory requirements, and promoting the widespread adoption of water reuse practices.

The availability of systems like VECOLI is completely aligned with the zero-pollution strategy: “To build a Healthy Planet for All, the European Green Deal calls for the EU to better monitor, report, prevent and remedy air, water, soil and consumer products pollution, among other things” (Commision s.f.) regarding water perspective.

The precision is large when the data used for training, accurately, describes the phenomenon that is predicted. The results provided by the virtual sensor offer a reliable tool for enabling the reuse of reclaimed water in slaughterhouses. This ensures the safe and continuous use of the water, significantly contributing to water savings while maintaining safety standards.

Keywords: Water reuse, safety, meat industry, drought, *E. coli* monitoring, water quality, aquaBio, virtual sensor.

REFERENCE

BETA, Grup. REAQUA Project. s.f. https://betatechcenter.com/es/proyectos/reaqua/.

Commision, European. «EU Action Plan: 'Towards Zero Pollution for Air, Water and Soil'.» s.f. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021DC0400&from=DA.