FEATURE ▶ Aeration

KEEPING UP WITH AERATION DEMAND

Integrating machine learning and bio-electrochemical sensors can increase energy savings and improve aeration control

Kash Delsoz Bahri, Rob Smith, Jack Ambler, Colin Ragush, Patrick Kiely

Water and wastewater treatment is typically among a community's largest energy consumers, accounting for 30% to 60% of energy use. It is no wonder that municipal utilities across the globe are seeking strategies for energy-efficient wastewater treatment.

Because aeration can account for the majority of the energy used for wastewater treatment — 50% to 75%, according to the 2021 manual *Energy in Water Resource Recovery Facilities* published by the Water Environment Federation (Alexandria, Virginia) — reducing energy demand for aeration is key to reducing costs and improving sustainability at water resource recovery facilities (WRRFs).

Aeration demand is dependent on a number of variables and can vary widely, requiring WRRFs to monitor continuously and make frequent adjustments. A convergence of two new technologies — bio-electrochemical sensors (BES) and machine learning (ML) — offers an opportunity to overcome the limitations of existing monitoring while improving the efficiency of wastewater aeration.


Paul Cockrell Photography

Aeration Considerations

Aeration demand is dependent on wastewater characteristics that vary daily, weekly, and seasonally. On top of that, transient discharges from industry and storm events add additional variability. As a result, aeration demand can vary by a factor of four or more, and frequent adjustments are required to optimize aeration. Continuous online monitoring of aeration demand is essential to achieve the goal of minimum energy use.

Estimation of aeration demand can be made by measuring oxygen demand. Biochemical oxygen demand (BOD) is a direct measurement of wastewater biodegradability, but it requires 5 days, which makes it unsuitable for process control. Measuring chemical oxygen demand (COD) is much faster. However, COD measures all oxidizable organics and thus is an indirect method of measuring oxygen demand that is biologically degradable. BOD and COD are accurate, standardized measurements of oxygen demand, but the limitations of the analytical methods mean that neither parameter is practical for aeration control.

Conventional aeration control typically is based on measurement of oxygen supply. Bacteria require oxygen to consume organics. Consequently, maintaining ample dissolved oxygen (DO) in a reactor ensures the reaction proceeds at a high rate. DO measurement is fast, simple, and accurate. However, it is a lagging indicator. Aeration control with DO is feedback control. The objective of feedback control with DO is to hold a DO concentration at some downstream location in the reactor that lags the introduction of the wastewater into the reactor. The location of the DO sensor(s) is critical to the stability of the control system.

Feed-forward control is superior to feedback control for responding to variable loading conditions and, as a result, has the potential to produce higher-quality effluent more consistently and at lower cost. Feed-forward control requires two components, measurement of the wastewater before it enters the reactor and a model to predict oxygen demand. Ammonia typically is the wastewater measurement as online ammonium measurement is reasonably reliable.

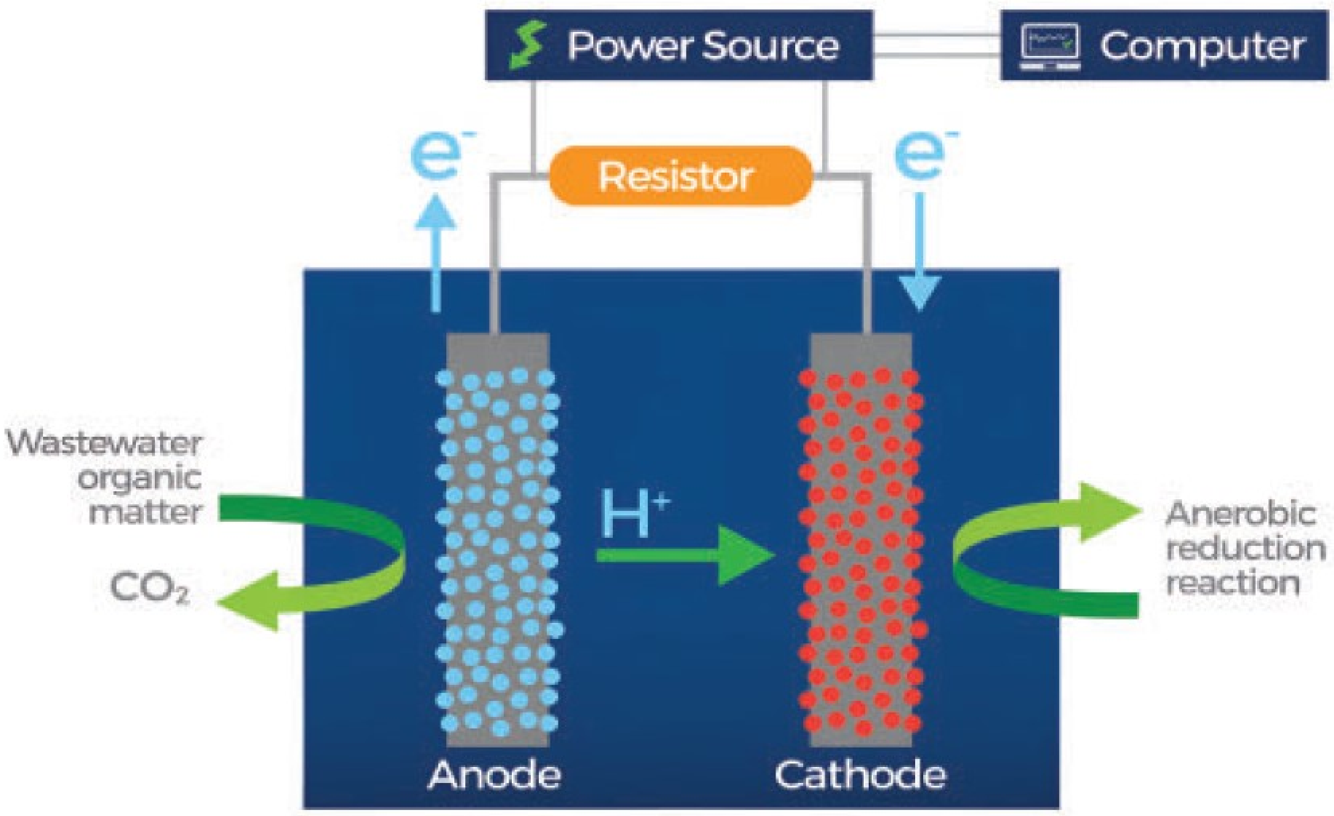
The complexity and accuracy of the mechanistic models typically used varies. As a result, another requirement of feed-forward control usually is a feedback component to correct for errors in the model. Direct measurement of biological oxygen demand with a BES and a more accurate model created with ML have the potential to improve feed-forward control and increase the efficiency of wastewater aeration.

A Living Sensor

A BES provides the capability to estimate wastewater BOD from front end of the wastewater treatment facility. This real-time reading at or near the influent location is a unique factor that makes it useful as part of organic loading prediction for improved aeration control accuracy.

The sensor consists of naturally occurring bacteria that form a biofilm attached to its conductive surface (see Figure 1, p. 42). As the biofilm consumes the wastewater, an electrical current (signal) is generated that provides a direct measure of oxygen demand by the in situ biology in real time. A higher signal means more readily consumable oxygen demand. As conditions in the wastewater change due to a slug load or stormwater, the signal increases or decreases according to those conditions. Essentially, the sensor provides a surrogate reading for the readily biodegradable COD or the food in the wastewater that is being consumed by this electrically connected biofilm.

Figure 1. BES Architecture



BES = bio-electrochemical sensor.
CO_2 = carbon dioxide.
H^+ = hydron.
e^- = electron.
e^+ = positron.

Biofilm will form on the anode side of the sensor and consume carbonaceous organics and respire electron as a result. The flow of the electron, measured by a control panel, indicates the biological activity and carbon consumption of the system.

A BES is simple to operate and maintain. Biofilms that cause fouling on traditional sensors are an essential component of the BES. Also, the sensor responds quickly and does not require manual, onsite calibration. Sensors installed at feed-forward locations of the wastewater treatment process — such as primary clarifiers and headworks — can be used to predict the aeration requirements dynamically.

The sensor output is a direct measure of consumable BOD (mg/L) being consumed in real time by the biomass in situ in the wastewater stream. If the sensor readings indicate challenging or problematic influent BOD conditions, it communicates data to operators as an alarm notification. The sensor signal can provide insights on the quantity of carbonaceous organic material entering the aeration tank and its effects on aeration demand.

Machine Learning

ML is a complementary technology that, coupled with a BES, can increase control accuracy further, resulting in more reliable treatment and greater energy efficiency.

ML, a subset of artificial intelligence, is dedicated to developing algorithms and models that emulate the human capacity for learning. These models are designed to learn from past data to detect patterns and correlations among various input parameters. Such learning enables them to forecast future behavior when they are presented with new and unseen data.

In a broader sense, an ML algorithm is a black box model of a system that links input data to output prediction without disclosing its internal workings. In the context of wastewater treatment, numerous variables — both quantifiable and beyond measurement — influence the demand for oxygen. ML offers a way to harness the available data by constructing a model that bridges the gap between observed variables and the predictive goal, thereby enhancing our understanding and management of these complex systems.

Typically, ML models undergo evaluation prior to their full deployment. This critical step using historical data sets sets aims to find which models most closely align with both the data and the specific problem statement. Statistically evaluating the outcomes of different models helps to identify the most effective model. Metrics used include R^2, a measure of model accuracy; and Root Mean Squared Error (RMSE), a measure of model precision.

This evaluation process involves dividing the data set into two segments: the training data set, on which the model is trained; and the validation/testing data set, which is used to evaluate the model's performance on data it has not seen previously. This approach ensures rigorous testing of the model's accuracy and precision. Once a model is adequately trained, it then can be applied to new data, which facilitates real-time prediction and decision-making.

Integrating BES Signal and Aeration Control Logic

Using a BES enables the optimization of aeration rates, allowing WRRFs to monitor the fluctuating patterns of organic load closely throughout the day and week and enabling immediate actionable data predictions. This approach aims to refine the efficiency of the aeration process and helps WRRFs adapt swiftly to the dynamic nature of wastewater treatment.

These sensors have demonstrated effectiveness in monitoring the biological nutrient removal process. As described in their 2023 *Environmental Science & Technology* paper, "Integrating Bio-Electrochemical Sensors and Machine Learning to Predict the Efficacy of Biological Nutrient Removal Processes at Water Resource Recovery Facilities," Emaminejad and colleagues developed models that predicted the amount of nitrate eliminated from the system, identifying the BES signal as a primary predictive factor. Encouraged by these outcomes, the authors applied similar models to aeration data sets to evaluate the capability of the BES to predict DO levels. Minimum data required to initiate developing an ML model are as follows.

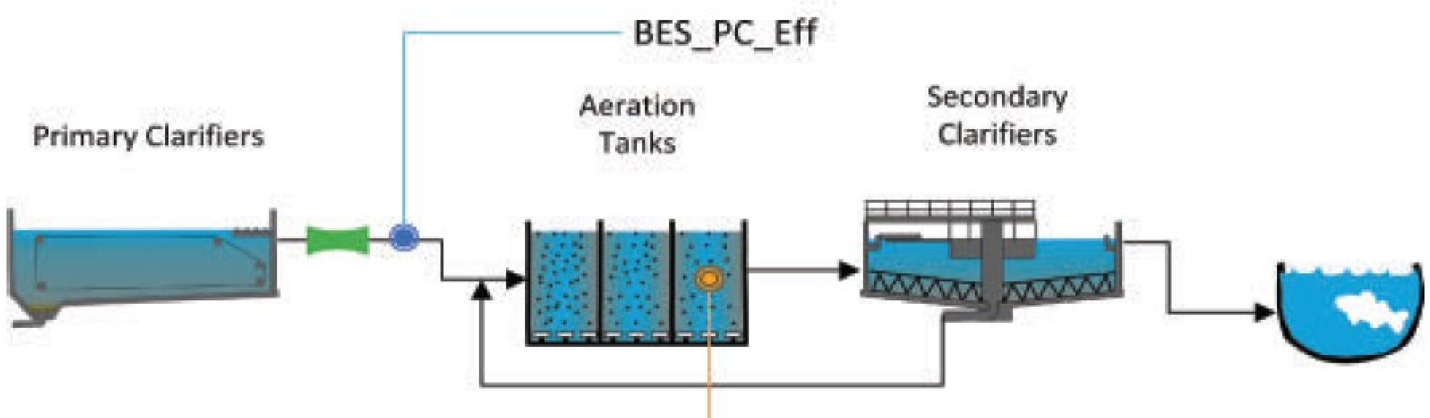
- Data interval should be at least hourly (every 10 to 15 minutes preferred).
- A minimum of 3 months of data is required to start the modeling process. (Working with less data is possible, but it will affect the development of the model and its accuracy.)
- DO data is an essential input, as it will be the target of prediction.
- Aeration rate also is mandatory, as it serves as the primary controlling variable for the control approach.
- Other process parameters such as influent flow rate, mixed liquor suspended solids (MLSS), and ammonia concentration can enhance the model's performance and provide insights into the intricacies of the system.
- There should be at least one BES at influent and/or aeration tanks.

The objective for developing the model is to create a tool capable of predicting oxygen demand and dissolved oxygen levels, minutes to hours ahead of time. By achieving this, the models either can guide operators to adjust the aeration process manually, or automatically supply a feed-forward element for the DO controller through model predictive control (MPC). This predictive capability aims to enhance the operational efficiency of aeration systems, ensuring optimal oxygen levels are maintained for the wastewater treatment process.

MPC employs a process model to forecast the outcome of the process based on inputs and disturbances. It uses an optimizer to adjust the controlled variables within the model, ensuring they satisfy various constraints applied to both inputs and outputs. Leveraging a model for prediction enables the controller to operate in a feed-forward manner, as it allows for more precise adjustments in response to real-time changes in the process.

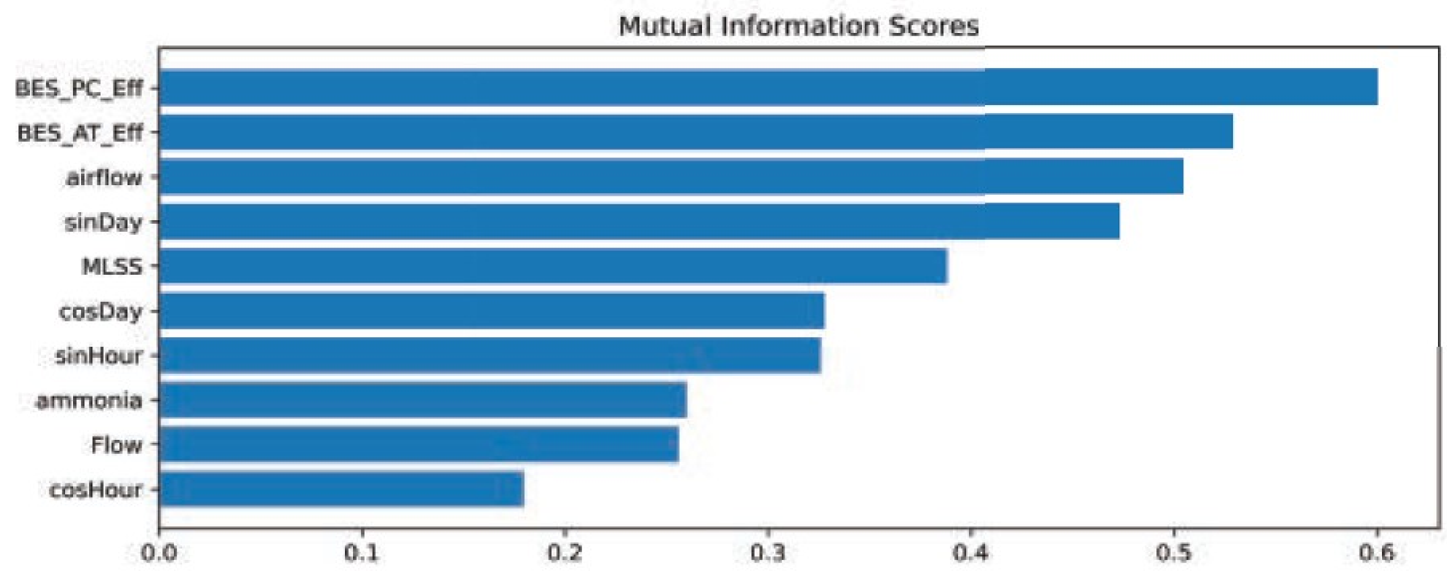
An advanced aeration control (AAC) strategy that incorporates MPCs uses three cascade control loops. The initial loop determines the ideal DO setpoints for each control zone, using the BES signal as a proxy for the combined BOD and nitrogenous oxygen demand. The second control loop focuses on computing the airflow required to maintain the DO levels at the target setpoints. In the third and final control loop, AAC specifies the valve position for every control zone, ensuring precise allocation of airflow to maintain the targeted aeration efficiency across the system. The second and third loops are the same as for conventional aeration control.

Figure 2. BES Configuration at WRRF



BES = bio-electrochemical sensor.
WRRF = water resource recovery facility.
BES_PC_Eff = sensor 1.
BES_AT_Eff = sensor 2.

Figure 3. Top Predictors at WRRF



BES_PC_Eff = sensor 1.
BES_AT_Eff = sensor 2.
sinDay = day (sine-encoded).
MLSS = mixed liquor suspended solids.
cosDay = day (cosine-encoded).
sinHour = hour (sine-encoded).
cosHour = hour (cosine-encoded.)

Test Case

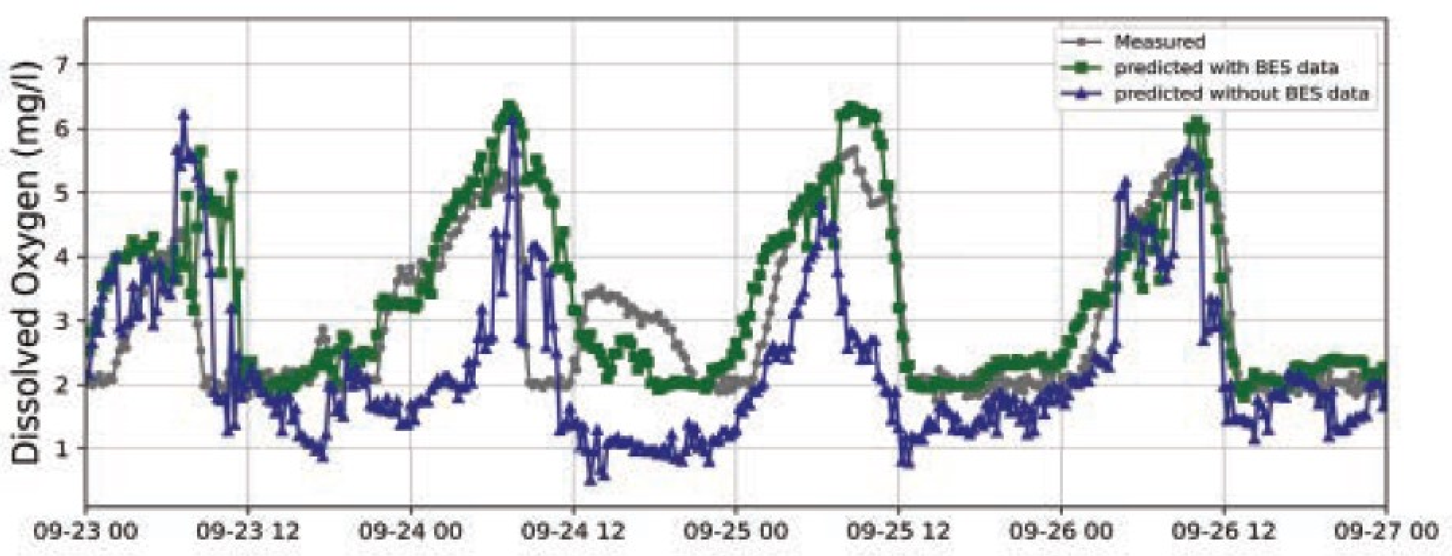
The team tested ML models that incorporate BES at a WRRF, using data collected May to September 2020 from a standard aeration unit. They installed two sensors — one after the primary clarifier and another at the effluent of the aeration tank (see Figure 2, p. 43). Other variables selected to develop the models included the aeration rate, flow rate, MLSS, ammonia concentration, and cyclical time variables reflecting diurnal and weekly patterns.

Before initializing the model development, the authors performed a series of analyses to determine the most desirable *features* — input variables — for the model. *Mutual information* (MI) — a measure of dependence between two variables — helped identified the top predictors. Based on MI results, BES signal has proven to be a strong predictor of DO. The effects of the remaining features are evaluated based on these results, and only the best features are included in the final model. Figure 3 (above) represents an example of ML performed on an aeration treatment case study. The BES signal, installed at the influent of the system, had the highest rank. This means that BES signal is the most strongly associated with the target variable being predicted. Similarly, other aeration case studies also showed the importance of the BES data in models.

Following data acquisition and cleaning, the team split the data set into training and test subsets to develop and test different models. This division was executed with a ratio of 70% for the training data set and 30% for the test data set. This methodology ensures that the model can generalize its learning to new, unseen situations, thereby providing a reliable measure of its predictive accuracy and robustness.

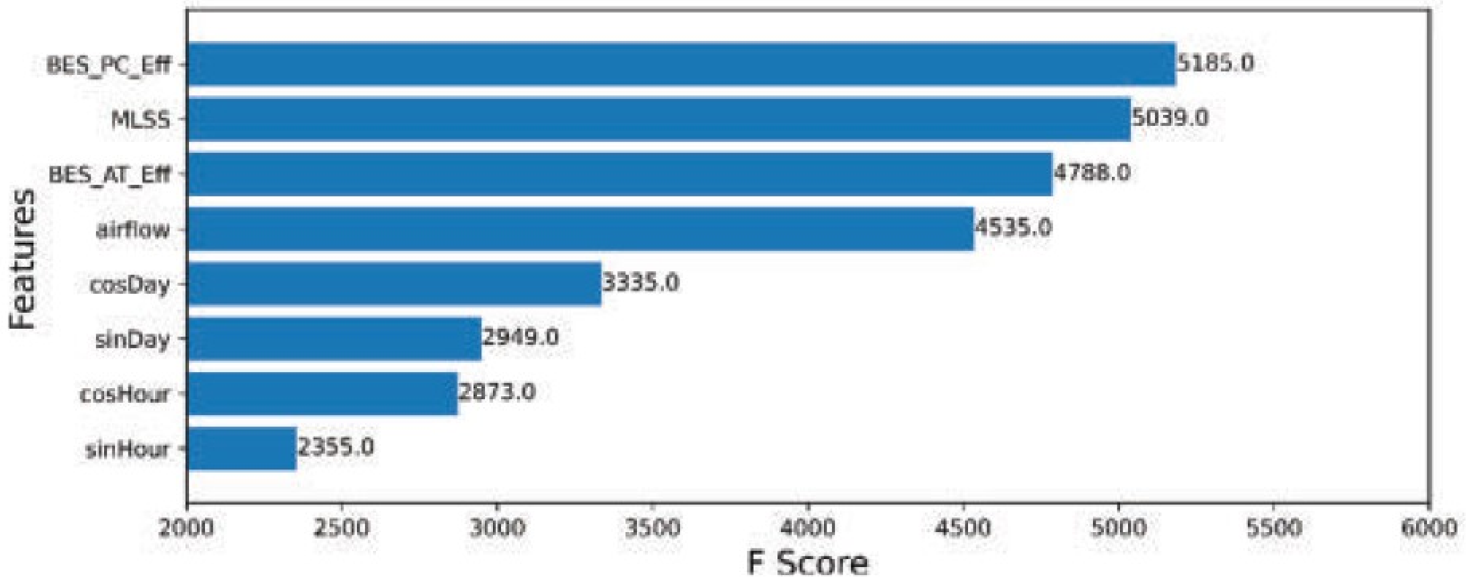
The team trained and tested different models including linear regression, Random Forest, and

Figure 4. Model Effectiveness With and Without BES Data



BES = bio-electrochemical sensor.

Figure 5. Feature Importance Analysis



BES_PC_Eff = sensor 1.
MLSS = mixed liquor suspended solids.
BES_AT_Eff = sensor 2.
cosDay = day (cosine-encoded).
sinDay = day (sine-encoded).
cosHour = hour (cosine-encoded).
sinHour = hour (sine-encoded).

XGBoost to identify the most effective model. Comparing model accuracies based on R^2 and RMSE metrics of the test data set helped identify the best model. The R^2 metric provides a measure of how well the model replicates observed outcomes, based on the proportion of total variation of outcomes explained by the model. On the other hand, RMSE offers a measure of the average magnitude of the errors between the values predicted by the model and the observed values, giving insight into the average prediction error. Higher R^2 and lower RMSE indicate a more accurate model.

The best model output — including the BES signals — resulted in an R^2 of 0.96 and RMSE of 0.35. This suggests that these models are effective in capturing the variance of the data set and making reliable predictions. However, when the sensors were excluded from the inputs, there was a decline in model performance; the R^2 decreased to 0.88, and the RMSE increased to 0.64 (see Figure 4, p. 44). This highlights the significant value of incorporating BES data into the model inputs to enhance the model's predictive accuracy and precision.

To assess the significance of each input parameter further, the team conducted a feature importance analysis. This analysis quantitatively evaluates the contribution of each input toward reducing prediction uncertainty, with higher scores indicating a greater impact on the model's predictions relative to other features. Based on the results, the BES signal from the primary clarifier was the top predictor, indicating the importance of the sensor in prediction (see Figure 5, above). This also suggests that a model with BES data as the feature likely can achieve higher accuracy with fewer required parameters as inputs for accurate prediction. More experiments with additional data sets are required to further evaluate this scenario.

Takeaways

The BES has great potential for wastewater treatment monitoring. It is a living sensor that provides an actual measurement of oxygen demand in real time. The sensors are simple to maintain, do not require calibration, and respond quickly to changing conditions. Furthermore, to function, the sensors rely on colonization by native bacteria — conditions that on conventional sensors would be considered fouling and interfere with proper function.

A machine learning model demonstrates the potential for BES in aeration control by predicting the DO concentration in an activated sludge aeration tank. The model inputs included routine measurements and signals from the sensors. The team found that sensor signals had the greatest effect on the DO prediction. Thus, the sensor signal showed itself to be a biologically relevant input that is essential for ML model accuracy.

***Kash Delsoz Bahri**, MASc, is a Lead Data Scientist and Client Manager; **Colin Ragush**, Ph.D., is a Senior Data Scientist; **Jack Ambler**, P.E., is Chief Technology Officer; and **Patrick Kiely** is Chief Executive Officer of SENTRY Water Technologies (Charlottetown, Prince Edward Island, Canada). **Rob Smith**, P.E., BCEE, Ph.D., is a senior process engineer with Black & Veatch (Overland Park, Kansas).*