Full-Scale Development and Piloting of a Hybrid Digital Twin for Wastewater Operations Optimization

B. R. Johnson[1]*, C. Yang[1], Leiv Rieger[2], J. Registe[3], K. Lesnik[4], A. Menniti[5], K. Villez[6], D. Pienta[7], H. Stewart[8], T. Johnson[9], Ivan Miletic[2] J. Rickerman[10], M. Oristian[11], T. Peterson[12]

[1]Jacobs, Greenwood Village, CO
[2]Jacobs, Toronto, ON
[3]Jacobs, Elizabeth, NJ
[4]Maia Analytica, Portland, OR
[5]Clean Water Services, Tigard, OR
[6]Oak Ridge National Laboratory, Oak Ridge, TN
[7]Jacobs, Seattle, WA
[8]Jacobs, Philadelphia, PA
[9]Jacobs, Charlotte, NC
[10]Jacobs, Wethersfield, CT
[11]Alexandria Renew Enterprises (AlexRenew), VA
[12]City of Tacoma, Tacoma, WA
*bruce.johnson2@jacobs.com

Full-Scale Development and Piloting of a Hybrid Digital Twin for Wastewater Operations Optimization

B. R. Johnson[1]*, C. Yang[1], Leiv Rieger[2], J. Registe[3], K. Lesnik[4], A. Menniti[5], K. Villez[6], D. Pienta[7], H. Stewart[8], T. Johnson[9], Ivan Miletic[2] J. Rickerman[10], M. Oristian[11], T. Peterson[12]

[1]Jacobs, Greenwood Village, CO
[2]Jacobs, Toronto, ON
[3]Jacobs, Elizabeth, NJ
[4]Maia Analytica, Portland, OR
[5]Clean Water Services, Tigard, OR
[6]Oak Ridge National Laboratory, Oak Ridge, TN
[7]Jacobs, Seattle, WA
[8]Jacobs, Philadelphia, PA
[9]Jacobs, Charlotte, NC
[10]Jacobs, Wethersfield, CT
[11]Alexandria Renew Enterprises (AlexRenew), VA
[12]City of Tacoma, Tacoma, WA
*bruce.johnson2@jacobs.com

Abstract
Full-scale piloting of a digital twin is being done as part of The Water Research Foundation project 5121: Development of Innovative Predictive Control Strategies for Nutrient Removal. The overall goal of this project is to develop and full-scale test a hybrid (machine learning + mechanistic model) nutrient management controller at four different water resource recovery facilities. The controller demonstrates both short-term optimization functions and longer-term predictive capabilities. This project also includes the design of the controller and performance at the WRRFs.

This paper provides results from three of the four case studies demonstrating the applicability and the benefits of the Hybrid Optimizer. An interesting result of all three pilots was the benefits of the additional information provided by the soft sensor/mechanistic model assisting with other operational aspects of the facilities.

Keywords
Digital Twin, Machine Learning, Hybrid Model. Full Scale, Soft Sensor

INTRODUCTION

This full-scale pilot of a machine learning based nutrient controller/digital twin is being done as part of The Water Research Foundation project 5121: Development of Innovative Predictive Control Strategies for Nutrient Removal. The overall goal of this project is to develop and full-scale test a hybrid (machine learning + mechanistic model) nutrient management controller at four different water resource recovery facilities (WRRFs). The controller demonstrates both short-term optimization functions and longer-term predictive capabilities. This project also includes the design of the controller, its performance at the WRRFs, and a description of the future improvements.

Advanced control adoption by utilities is one major industry trend. Another is the high level of interest and research into the benefits and applicability of machine learning approaches to approaches to help optimize systems throughout our society. Machine Learning (ML) systems do not rely upon mechanistic predictions but find empirical relationships which can be used to deal with uncertainty and variability in a way that conventional controllers cannot capture. Importantly, machine learning models are designed for fast calibration with online sensor data of high frequency, thus enabling quick adaptation to changes that are difficult to predict through conventional models.

Today, the use of machine learning in WRRFs nutrient management is in its infancy with very limited full-scale testing.

The hybrid digital twin developed on this project has been named ODIN (Operational Decision-making Information Network). It is important to note that currently the control for all four pilots is advisory only and that ODIN has no direct control authority over plant operations. The four full scale pilots are:

1. Clean Water Services Durham AWTF (CWS): ODIN is being used to recommend primary clarifier alum addition setpoints based on forecasted phosphorus loads and target bioreactor ortho-phosphate mass loading targets with the goal of minimizing chemical usage.
2. Agua Nueva WRF (ANWRF): ODIN is being used to recommend dissolved oxygen (DO) setpoints twice a day for the bioreactor systems as well as wasting rates with the goal of minimizing overall energy usage at the facility.
3. AlexRenew: ODIN is recommending equalization discharge and return pumping setpoints with the goal of minimizing overall energy and chemical usage at the facility, while maintaining a very low effluent total nitrogen (TN).
4. City of Tacoma North End Plant (Tacoma): ODIN is being used to recommend primary clarifier alum and polymer addition setpoints based on forecasted flows and loads while achieving target downstream trickling filter performance with the goal of minimizing chemical usage.

METHODS

The Hybrid Optimizer applies a hybrid digital twin approach combining mechanistic and data-driven models to: a) enable a plant-wide set of soft sensors based on a minimum number of required physical sensors. This makes digital twin solutions affordable for small to mid-sized plants; b) provide operational recommendations based on optimization results driven by emulator models and/or optimization algorithms; c) provide performance monitoring modules such as sensor or controller performance; and d) enable advanced users to execute data analytics with pre-configured tools.

The mechanistic model is the core of the Hybrid Optimizer, establishing the foundation of the digital twin model and enabling soft sensors to be developed for the entire facility. The soft sensor results (synthetic data) sequentially pass through three data-driven models (forecaster, emulator, and optimizer) to provide operational recommendations to the operating staff at a facility. The Hybrid Optimizer automates the calibration process of the mechanistic model by adjusting model parameters to match model results to the measured sensor or laboratory data.

The first step of the Hybrid Optimizer concept is to estimate dynamic raw influent concentration profiles that match the observed airflow rates and dissolved oxygen (DO) concentrations in the bioreactors. These airflow rates are a good indicator of the influent dynamics (e. g. COD, ammonia) if dissolved oxygen concentrations are known. The variation of total air delivered to the bioreactors results from the variation of pollutant loadings (flows and concentrations) into the plant, and thus can be used to estimate the influent loading profiles. The impact of other parameters such as sludge retention time and operational settings are covered by the mechanistic model.

Hybrid Optimizer Workflow

The Hybrid Optimizer is a cloud-based controller that consists of seven major components. Figure 1 shows a diagram of the Hybrid Optimizer workflow:

1. Data Ingestion: Cloud-based automatic tools to build live data stream connection with the Hybrid Optimizer regarding sensor and laboratory data.
2. Data Cleaning and Imputation: These tools clean and impute (infill) the raw data to provide a reliable and coherent dataset for the downstream mechanistic model and forecasting tools.

3. Soft Sensor: A soft sensor is an algorithm that combines signals from various available physical sensors to derive the current or future values of a variable of interest which is not or rarely measured, or not measured yet at the time of estimation. It reduces the number of physical sensors needed for advanced nutrient control or can be used to validate sensor measurements. For example, an ammonia soft sensor could be used to auto-correct or temporarily replace an existing ammonia sensor with poor performance. The Hybrid Optimizer employs an auto-calibrated mechanistic model as its central component to estimate the 15-minute interval concentrations of raw sewage influent as a multi variable soft sensor (Yang et al., 2023). The Hybrid Optimizer can use any process simulator, but the WRF project integrates the SUMO™ (www.dynamita.com) simulator as the base for its mechanistic process model.
4. Influent Forecaster: A machine learning tool to predict both influent flow and concentrations over the next 24 hours at 15-minute intervals. Its inputs include soft sensor results, historically measured flows, and internet-based 24-hour weather forecasts and river water levels. Once the influent forecast is generated, it is then used in the calibrated mechanistic model to forecast overall facility performance.
5. Model Emulator: A machine learning model provides high-fidelity approximations of the mathematical relationships contained in the mechanistic model. This emulator model helps to speed up the optimal solution-searching process of the optimizer.
6. Optimizer: The optimizer utilizes machine learning and/or optimization algorithms to perform multi-objective optimization to determine the optimal operational settings that align with the control goals.
7. User-Interface: A multi-level user interface (UI) is being customized for each pilot site including the three levels: 1) Operator, 2) Analyst, and 3) Engineer. The goal of the Operator UI is to provide a simple, user-friendly front end that operators can quickly reference for recommended settings. The Analyst UI provides information to analyse controller performance, and the Engineer UI allows for the evaluation of the different models and overall system performance including data quality control.

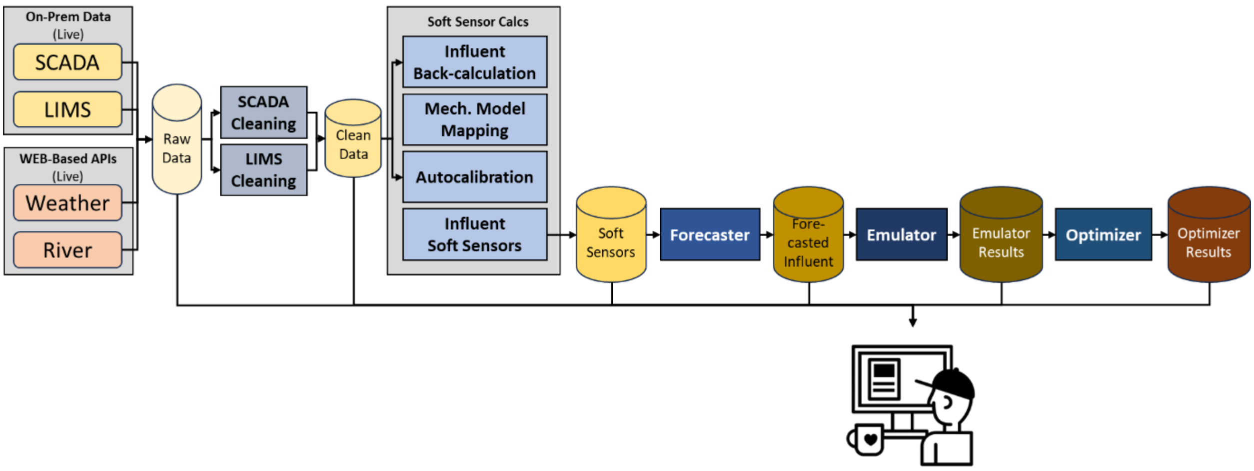

Figure 1: Hybrid Optimizer workflow.

A key functionality in the Hybrid Optimizer is the Soft Sensor. Using the actual measured flows and the mechanistic model, this tool back calculates what dynamic 15-minutes concentrations of the COD, TKN, and TP must have come into the facility to produce the observed dissolved oxygen concentrations (DO's) at the given air rates (Yang et al., 2023). In doing these calculations, the soft sensor also estimates all the other wastewater concentrations throughout the process model, thus providing a complete digital twin of the modelled parts of the facility.

One of the most important parts of a digital tool like the Hybrid Optimizer is to provide targeted information for the users. The philosophy behind the Hybrid optimizer user interface is to make this tool easy to use for operations staff (Menniti et al., 2023). To accomplish this objective, the focus was put on two interface design goals:

1. Identify user priorities and mind-set at every step.
2. Any new operational tool must make the life of the front-line operator better in some way to maximize adoption.

To guarantee accurate operation of the Hybrid Optimizer, a quality control concept has been developed. The following list describes a set of plots providing visual support to judge soft sensor performance including statistically sound criteria to trigger *warnings* (to enable operating staff to take corrective actions) and *alarms* (soft sensor values should not be used in applications). Example QA/QC plots are provided in Figure 5 and Figure 7 for Clean Water Services.

1. Control Chart on metric: A calculated performance metric (t-score, Milton and Arnold, 1995) is fed to a control chart using a warning limit based on a 95% confidence interval. The alarm limit is set to a confidence interval of 99%. A warning is triggering corrective actions, whereas an alarm flags a soft sensor value as not usable in any application. The main point here is that the metric is independent from the concentration or flow level and therefore robust against low number bias. The control chart is set up to trigger early warnings (using the warning limits) and to provide criteria to decide when the model cannot be used for recommendations (the alarm limits).
2. Process Plot: Modelled and measured variables plotted as time series. This gives a better view to eyeball performance, but it is difficult to identify drifts within high resolution data (therefore the error plot).
3. Unscaled Error Plot (Error = modelled – measured): This gives an idea of the actual value of the model error to judge if the error is acceptable.
4. x-y Plot: This is to assess the overall model performance. The plotted prediction intervals provide statistics but are not meant to be used as criteria for warning and alarming (this would overlap with the metric used in the control chart).

RESULTS AND DISCUSSION

This section provides results from three full-scale Hybrid Optimizer deployments.

Clean Water Services (CWS) - Durham WRRF

Objectives

The primary objective at Durham WRRF is to a) provide a soft sensor for the primary effluent phosphorus load with the ability to forecast 24 hours in advance, and b) optimize the primary alum dosing system by recommending chemical dosage rates. Acceptance criteria are focused on the achievable accuracy of the soft sensor and forecaster, which should consistently stay within a 20% error margin. Secondary objectives include:

- Full-scale assessment of accuracy for 24-hour forecasted predictions of:
    - Influent flow
    - Influent and primary effluent ammonia loading
    - Influent and primary effluent soluble COD loading
- Full-scale assessment of the Hybrid Optimizer data cleaning pipeline and analysis of the discovery time of erroneous data

- Assess extending the Hybrid Optimizer to additional applications such as DO control and providing information on the health of the biological phosphorus removal process.

Hybrid Optimizer Setup

The mechanistic process model used for Durham WRRF includes the liquids train through secondary effluent as shown in Figure 3. This includes alum dosing, primary clarification, secondary biological treatment (A_2O) with both nitrogen and phosphorus removal (aeration basins and secondary clarifiers), and plant recycles. Only one of the four bioreactor trains is included to estimate diurnal influent strength changes.

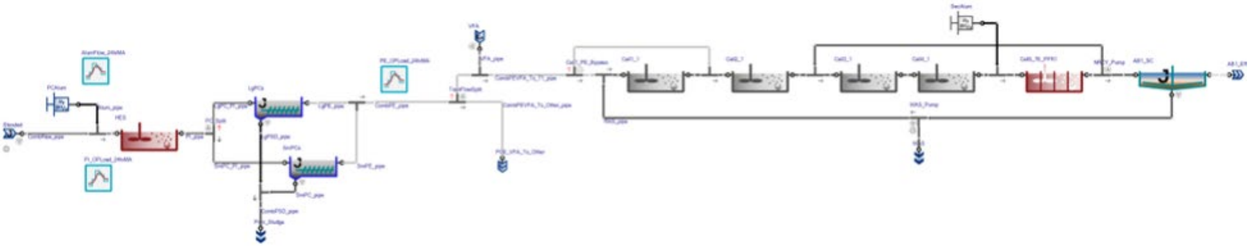

Figure 2: Clean Water Services - Durham WRRF mechanistic model process flow diagram .

Results

The following section shows results for the ortho-phosphate soft sensors as the main objective, the applied quality control, and the developed dashboards. The model prediction accuracy of the Hybrid Optimizer has been continuously improved, and the following results are derived from the latest model generation in a test environment rerunning all the data collected for the time frame shown.

Soft Sensor: Figure 3 provides soft sensor estimated profiles over 13 months (January 2023 to February 2024). These results were calculated on historical data but were done ignoring data that was not available at the time of the estimation. This was done to provide a valid approximation of live operation scenario. The first plot is of primary influent ammonia. The second and third graphs show ortho-phosphate in the primary influent and primary effluent, respectively, which are the target variables for chemical dosing control. As shown in the graphs, the soft sensor profiles are visually reasonable and likely adequate for real-time control especially considering no dynamic phosphorus data is available.

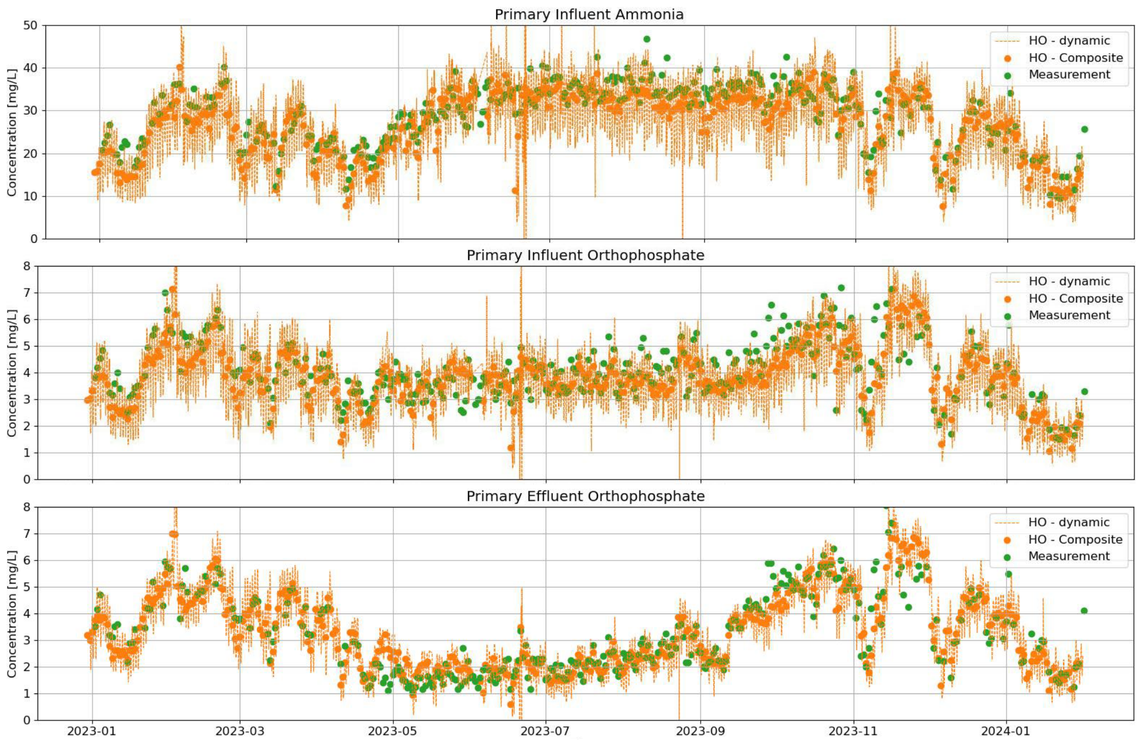

Figure 3: Soft sensor results for Durham WRRF: primary influent ammonia (top), primary influent ortho-phosphates (middle), and primary effluent ortho-phosphates (bottom).

Forecaster: This model provides 24-hour forecasts to enable facility staff to take actions before a higher load hits the plant. Figure 4 shows forecaster results for ortho-phosphate in the primary effluent between January 2023 and February 2024 and compares with the soft sensor results (Sumo), and the actual measurements. Like the previous graphs, all forecasts were done without foreknowledge of current laboratory or actual performance during the forecasted period. The forecast model was regularly retrained every three months on the previous three months of results to show the improvements in the predictions with a growing database. Going forward the goal is to add an additional check based on a goodness of fit metric when the measured data is available. The forecaster reliably predicts the trends of ortho-phosphate in the primary effluent.

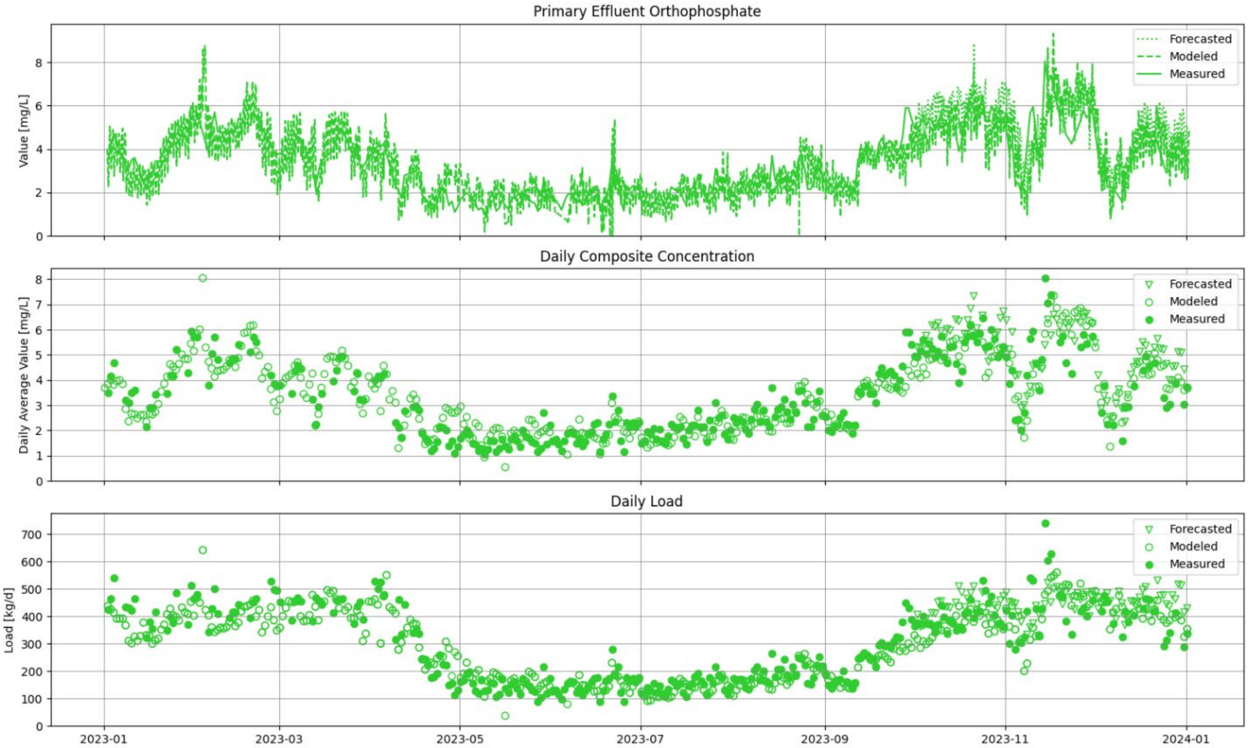

Figure 4: Forecaster ortho-phosphate results for Durham WRRF: primary effluent diurnal variations (top), comparison between measured, simulated, and forecasted composite samples (middle), and comparison of measured, simulated, and forecasted load.

Quality Control: Figures 5, 6 and 7 show the quality control setup for the main soft sensors (ortho-phosphate in the primary effluent). The quality control concept will be applied after the accuracy of the soft sensor has been accepted as sufficient. The control chart (Figure 5, bottom) shows larger variations in the metric between January and April 2023 and again from December 2023 to January 2024. The reason is that during May and September, the plant must fulfil stricter effluent limits and therefore doses alum. During this phase, as shown in Figure 6, the soft sensor uses an additional auto-calibration routine improving the prediction accuracy. If the quality control system would have been running live, the following notifications should have been triggered: In February 2023 and again in November 2023 warning limits are hit. Alarms were flagged in October and November 2023. The middle plot shows that overall, the soft sensor follows the measured trends very closely. The bottom plot confirms the control chart metric but gives a better idea to operations staff on the extent of the error between modelled and measured values.

Figure 5 shows a comparison between a) measured and modelled and b) measured and forecasted ortho-phosphate in the primary effluent in form of error plots over time. The top plot is displaying the raw error for the soft sensor (modelled) with a Mean Absolute Error (MAE) of 0.49 ± 0.42 mg

P/L and for the forecaster (starting in October 2023) with a Mean Absolute Error (MAE) of 0.91 ± 0.61 mg P/L.

The bottom plot uses percent error and results in a Mean Absolute Percentage Error (MAPE) of 17.95 ± 18.46 % for the soft sensor, and 22.16 ± 18.19 % for the forecaster.

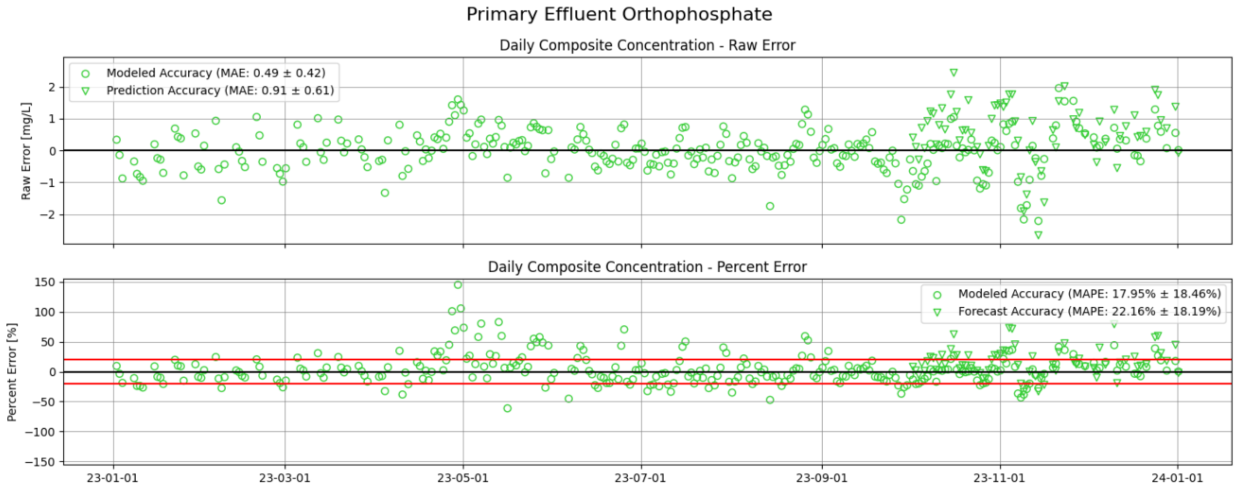

Figure 5: Error plots for modelled and forecasted ortho-phosphate compared to the measured composite values. Top plot: Raw error in mg P/L and bottom plot: Mean Absolute Percentage Error (MAPE) in percent. The forecasted results only start October 2023. The red band indicates the targeted error margin.

Figure 6 shows the same plots as Figure 5 for only the nutrient removal season between May and September with chemical phosphorus removal and an additional auto-calibration routine. The accuracy for the soft sensor is slightly lower for the raw error with a Mean Absolute Error (MAE) of 0.41 ± 0.35 mg P/L, but the percent error is slightly higher with a Mean Absolute Percentage Error (MAPE) of 19.28 ± 17.66 %.

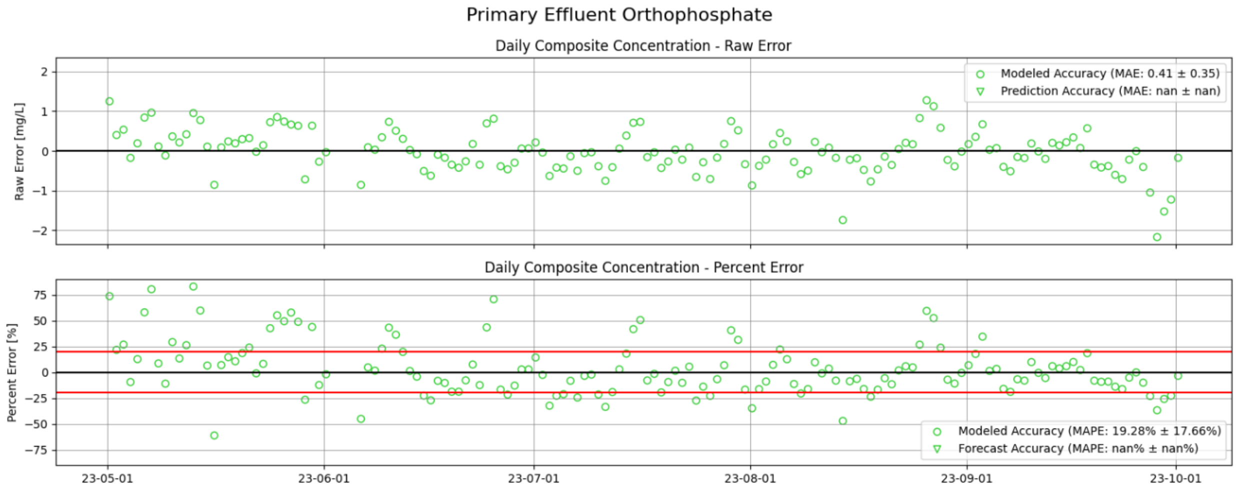

Figure 6: Error plots for the nutrient removal season (May to September) with chemical phosphorus removal. An additional auto-calibration routine was deployed during this time.

Figure 7 shows an x-y plot providing a visual impression of the overall trueness and precision of the soft sensor compared to measured data. The 95% and 99% prediction intervals are used to put new data into context. In this case the intervals were calculated based on the full data set, but in the on-line application it is envisioned to define a period with acceptable soft sensor performance and compare new data against this baseline performance. Criteria for retuning or retraining will rely on an input data check to ensure that the model conforms well to current normal plant operations.

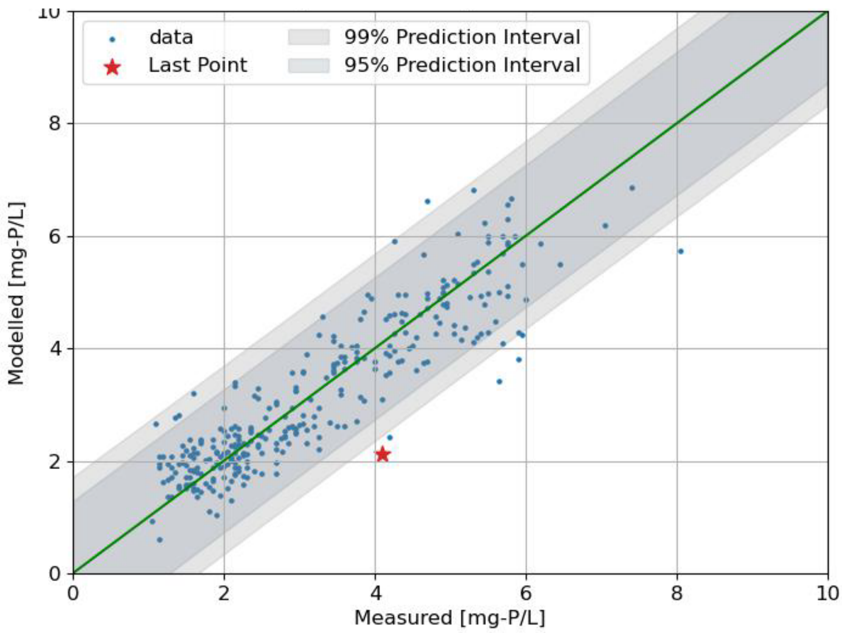

Figure 7: x-y plot of the primary effluent ortho-phosphate to analyse soft sensor performance at Durham WRRF.

Dashboard: This section describes the status of the three dashboards discussed earlier for deployment at the Durham WRRF.

- *Operator Dashboard*: Quick recommendations

  For the Durham WRRF, the Hybrid Optimizer recommendations are not connected directly to controllers and actuators for cybersecurity reasons. Therefore, the hybrid Optimizer recommendations are manually entered into SCADA menus at fixed times of the day. This step-by-step approach helps build the trust of the staff by integrating them in decision making to accept or reject the recommendations.

  The operator dashboard (Figure 8) includes only actionable information and is meant to be used to understand and rapidly implement the recommended control set point(s) for the day. This information could also be provided with daily emails, as previous experience at the Agua Nueva project site suggests that recommendations are likely to be consumed by the email or text communication rather than from the dashboard. Figure 8 shows the operator dashboard with the recommended alum dosing.

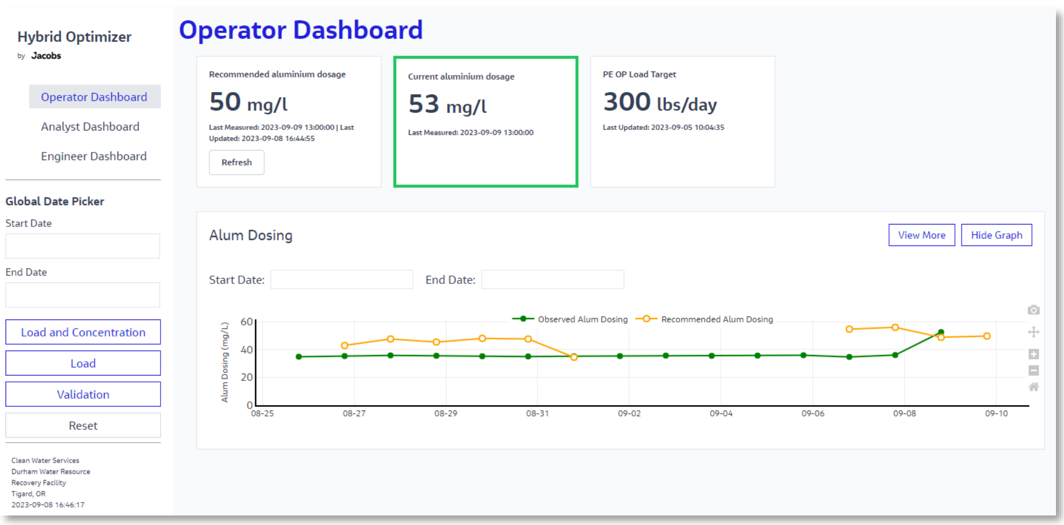

Figure 8: Hybrid Optimizer Operator dashboard at Durham WRRF.

- *Analyst Dashboard*: Performance review

  The analyst dashboard is designed for quality control and assurance purposes, ensuring the Hybrid Optimizer is generating rational information and that the recommendations are valid. The journey mapping process used to develop the UI approach highlighted the importance of including trust building elements into the dashboard (Figure 9). A simple control chart has been included to aid in this purpose. This control chart has control limits to benchmark the expected level of accuracy of the system and demonstrates if the hybrid optimizer is predicting with acceptable accuracy over time.

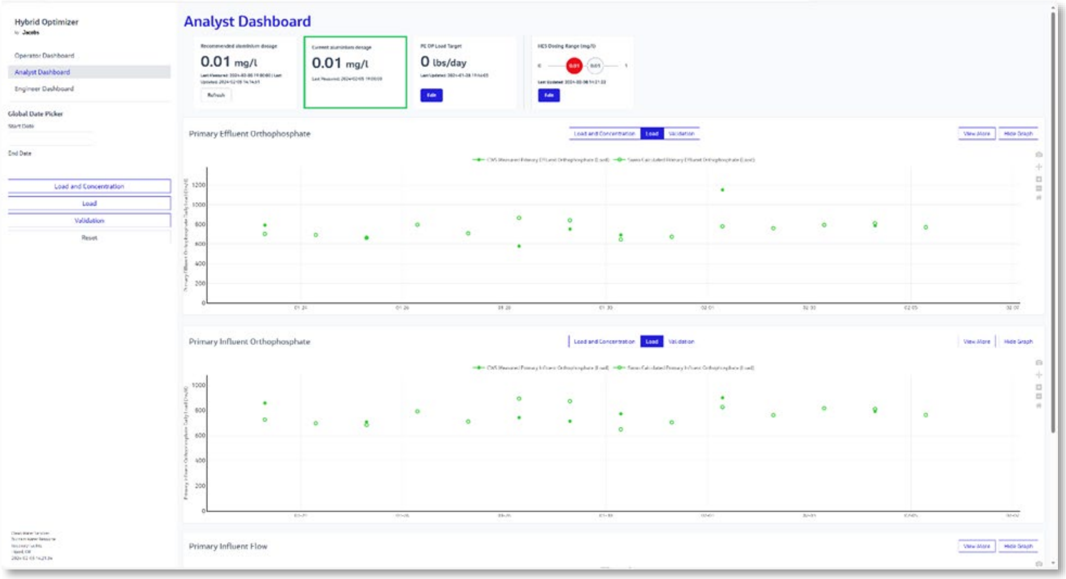

Figure 9: Hybrid Optimizer Analyst dashboard at Durham WRRF.

- *Engineer Dashboard*: Detailed Analysis

  The intention of the engineer dashboard (Figure 10) is to make inner workings of the Hybrid Optimizer models accessible to advanced users. Multiple models collaborate seamlessly to execute the Hybrid Optimizer control functionality. This dashboard is intended as a first stop in troubleshooting if the accuracy of the system degrades and in understanding the wider operation of the system.

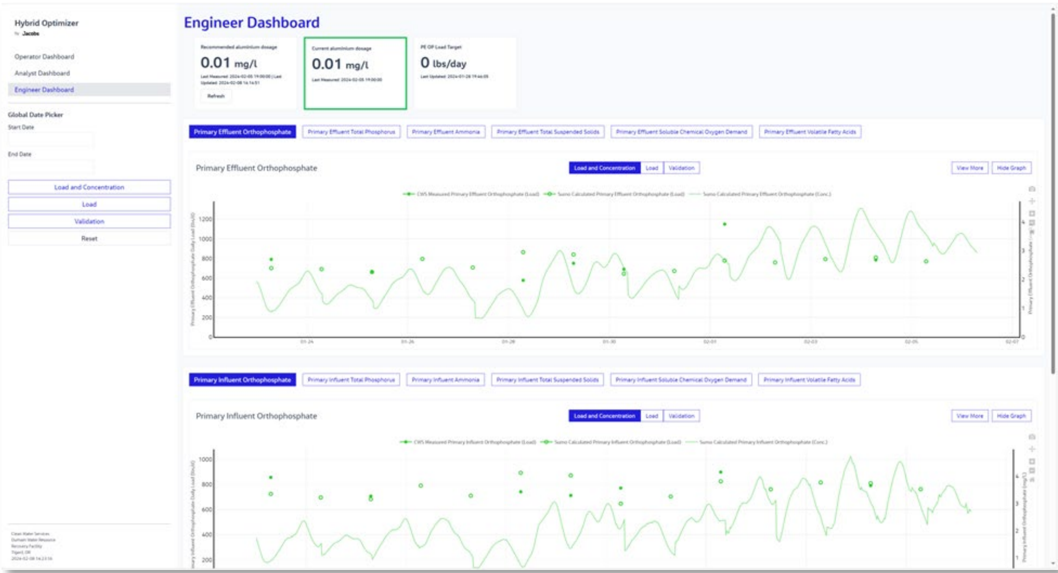

Figure 10: Hybrid Optimizer Engineer dashboard at Durham WRRF.

Discussion and Lessons Learned

The main objective of the Durham WRRF installation is to provide actionable insights into primary effluent ortho-phosphate profiles to support operation in managing primary effluent phosphorus loads. The Hybrid Optimizer has been collecting live data for more than one year enabling continuous refinements to the models. Moving from sporadic lab measurements of 24-hour composite samples to 15 min soft sensor values has the potential for enabling a better understanding of the process dynamics and will support the decision process to accept or reject the alum dosing recommendations by the Hybrid Optimizer. An additional benefit from the Hybrid Optimizer of nearly equal value to the staff, is the calculation of dynamic wastewater concentrations such as ammonia and COD in the primary influent and effluent.

Essential to the adoption of the new tool by plant staff was to involve them early in the design process and customize the three dashboard levels to provide tailored information according to their job role. The next step will be to roll out the dashboards to the plant staff and refine in regular feedback sessions.

The Hybrid Optimizer runs through a complex workflow and faults could happen at every step. To mitigate the risk of wrong conclusions or recommendations, an extensive data cleaning and imputation pipeline has been developed and a novel quality control concept allows early detection if one model drifts away. A next step is to automate the quality control concept and integrate in an alarming system to inform the user and enable corrective actions.

Pima County - Agua Nueva WRF

Objectives

The objective at Agua Nueva is to optimize the aeration system to reduce aeration costs. Bioreactor 4 will receive daily recommendations for the DO setpoint.

Hybrid Optimizer Setup

The mechanistic process model of the Agua Nueva WRF includes the liquids train through secondary effluent as shown in Figure 11. This includes ferric dosing, primary clarification, secondary step feed biological treatment for ammonia and BOD removal, secondary clarifiers, and plant recycles. Only one of the four parallel trains is modelled.

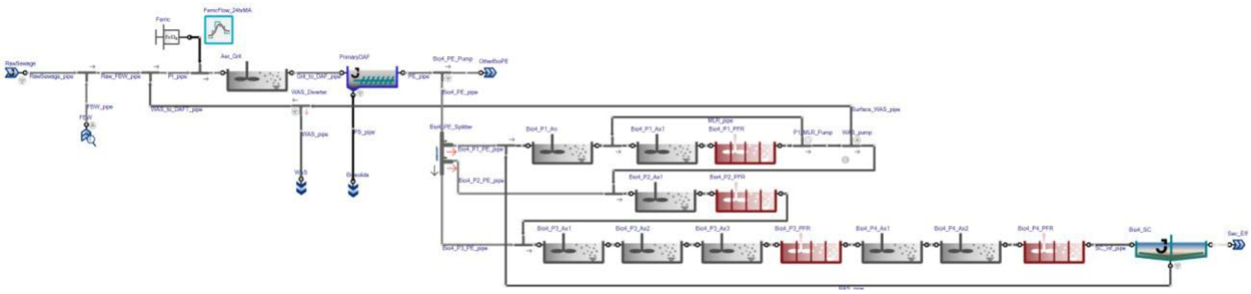

Figure 11: Pima County - Agua Nueva WRF mechanistic model process flow diagram as implemented in SUMO.

Results

The Hybrid Optimizer digital twin (soft sensors) implementation has been tested successfully for over 4 months. Figure 12 shows a comparison between primary influent measured laboratory data (24-hour composite samples) and synthetic composite data (derived from the simulated dynamic data to enable a direct comparison). The Hybrid Optimizer predictions follow the measured data closely.

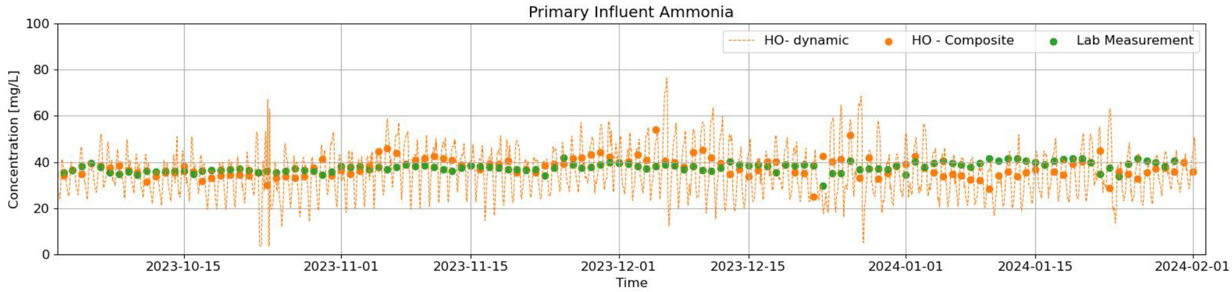

Figure 12: Comparison between five months of measured lab data (green dots) and simulated composite data (orange dots) calculated from the dynamic data (orange dashed line) for Agua Nueva WRF.

Figure 13 demonstrates the excellent fit between measured and simulated airflow rates. The total airflow is used as the target variable in the back-calculation of the influent ammonia load by the soft sensor.

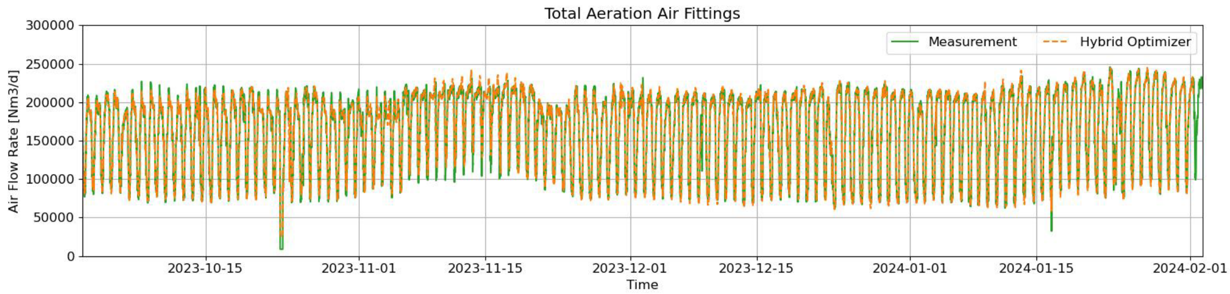

Figure 13: Comparison between measured (green) and simulated (orange) total airflow rates for Agua Nueva WRF. The fit results from back-calculating the influent ammonia concentrations to match the measured airflow rates by the soft sensor.

Figure 14 shows a good fit between the measured and modelled primary effluent total suspended solids (TSS) concentrations. However, Figure 15 shows that during October and November 2023 the Hybrid Optimizer was not matching the measured mixed liquor suspended solids (MLSS) concentrations. Upon investigation it was found that the system incorrectly calculated the wasting rate when Bioreactor 3 (which is not part of the digital twin model) was taken out of service. Bioreactors 3 and 4 share a wasting pump and thus the wasting rate was cut in half for modelling purposes. When Bioreactor 3 was out of service, the model falsely received wasting rates half of what they were, thus driving up the MLSS concentrations. It can be seen from the MLSS graph, that once Bioreactor 3 went back into service, the MLSS levels were brought back into relatively good agreement.

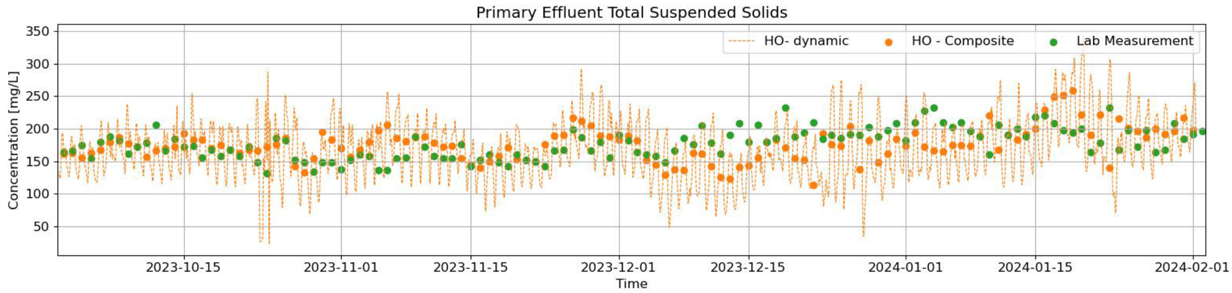

Figure 14: Comparison between measured (green) and simulated (orange) primary effluent TSS concentrations at the Agua Nueva WRF.

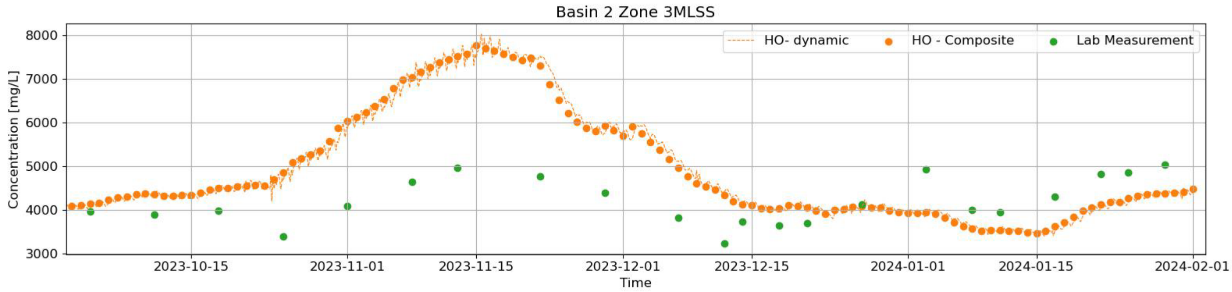

Figure 15: Comparison between measured (green) and simulated (orange) MLSS concentrations at the Agua Nueva WRF.

One finding from the Hybrid Optimizer was that the plant's calibrated ammonia PE probe was always measuring high as compared to the soft sensor back calculation (Figure 16). A detailed analysis found that the plant staff were calibrating the probe based on test kit results. However, the used test kits showed a sensitivity to the high sulfides concentrations at the plant resulting in a positive bias. After a sulfide inhibitor was added to the laboratory method, the results between sensor and soft sensor were much closer.

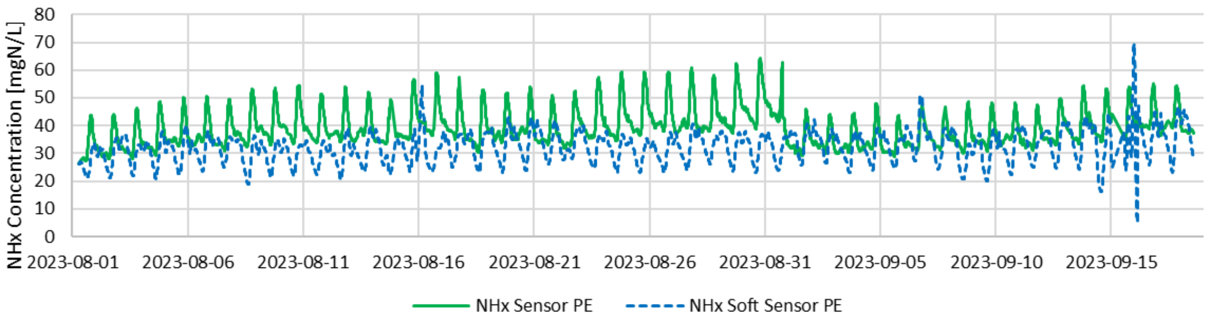

Figure 16: Comparison between the physical and soft sensors for NHx in the primary effluent of the Agua Nueva WRF.

Discussion and Lessons Learned

While the primary goal of the Hybrid Optimizer was to help reduce aeration demands, the application of this tool has resulted in several additional insights related to instrument reliability, improved DO control and effluent quality. One of the primary lessons learned is that a digital twin needs robust detection and knowledge of units being out of service, even if it is not directly part of the digital twin, as illustrated in the mixed liquor results of the model. The identification of the biased laboratory measurements highlights the benefits of using a mechanistic model, which is based on mass balances.

AlexRenew WRRF

Objectives

The objective at AlexRenew is to optimize the primary effluent equalization via an existing equalization tank (Nutrient Management Facility, NMF) flow controls in order to minimize aeration and supplemental carbon usage. In advisory mode, the Hybrid Optimizer is making recommendations for the NMF flow control setpoints three times a day. By stabilizing the nitrogen load to the bioreactor, staff can then select more optimized setpoints for both DO control and methanol addition, thus reducing overall operations costs.

Hybrid Optimizer Setup

The mechanistic process model for AlexRenew is a whole plant model that includes the full liquids train and a simplified version of the solids train as shown in Figure 19. This includes primary clarification, primary effluent equalization in the NMF, secondary step-feed biological treatment for total nitrogen and chemical phosphorus removal (aeration basins, supplemental carbon dosing, ferric dosing, and secondary clarifiers), tertiary treatment, thickening, digestion, and dewatering.

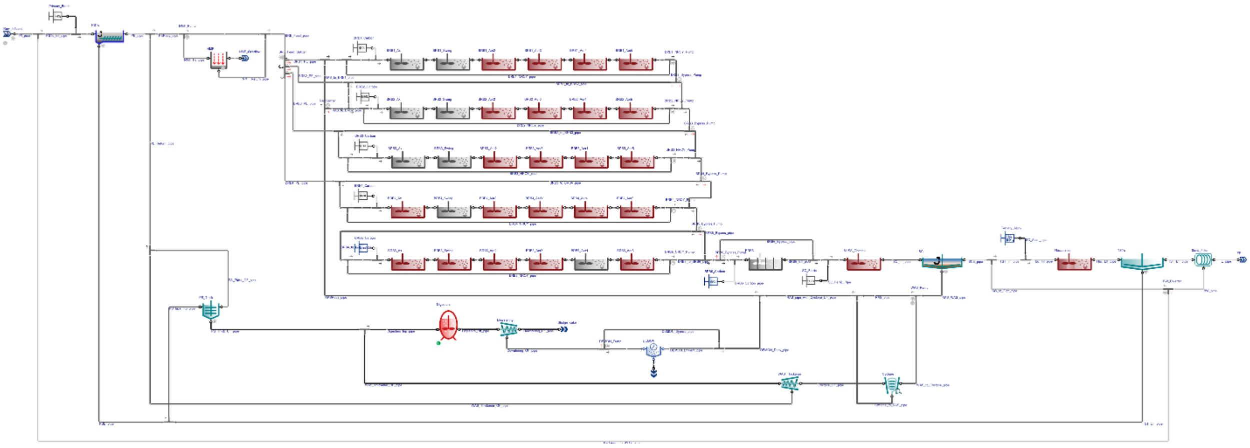

Figure 17: AlexRenew WRRF whole-plant mechanistic model process flow diagram as implemented in SUMO.

Results

The digital twin in the Hybrid Optimizer has been successfully operated at AlexRenew WRRF for over 4 months, while the recommendation system has been deployed for 1 month. Figure 20 shows that the Hybrid Optimizer influent soft sensor for ammonia closely follows the laboratory measurements. The larger difference at the end (greyed period) is discussed below.

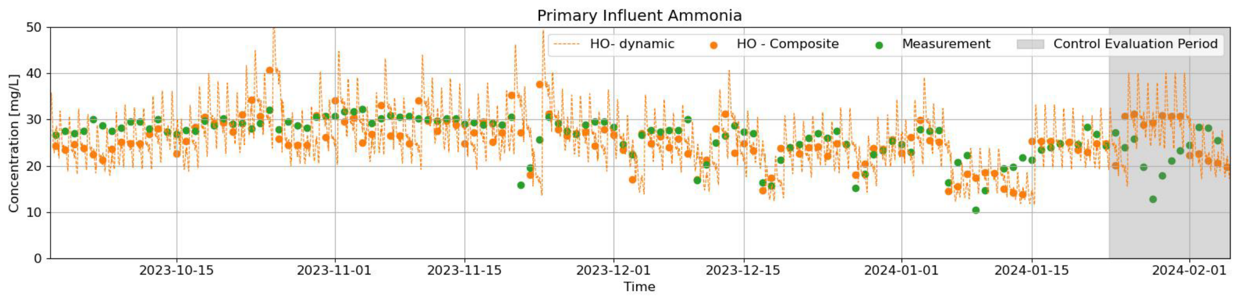

Figure 18: Comparison of measured (green) and simulated (orange) primary influent ammonia at AlexRenew WRRF.

At the AlexRenew Hybrid Optimizer implementation, special attention was given to the flow forecasting model to provide future flow information to the optimizer and make the best use of the Nutrient Management Facility (NMF). The developed forecaster uses historic flow data, day of the week and day of the year information, weather forecasts, and water level and flow information from the nearby river (Cameron Run).

Figure 19 (top) shows the measured (green) and forecasted (orange) influent flow for the last two months. The bottom graph provides the forecasted and measured rainfall. The large spike in forecasted flow on January 26th can be linked to a high forecasted precipitation, which did not happen. The extreme rain event on January 28th is not fully captured by the forecaster although the rain forecast was very accurate. On the other hand, there are potential faults in the measured data

(January 22nd and 31st). The high variation led to inaccuracies of the forecast model on the next day.

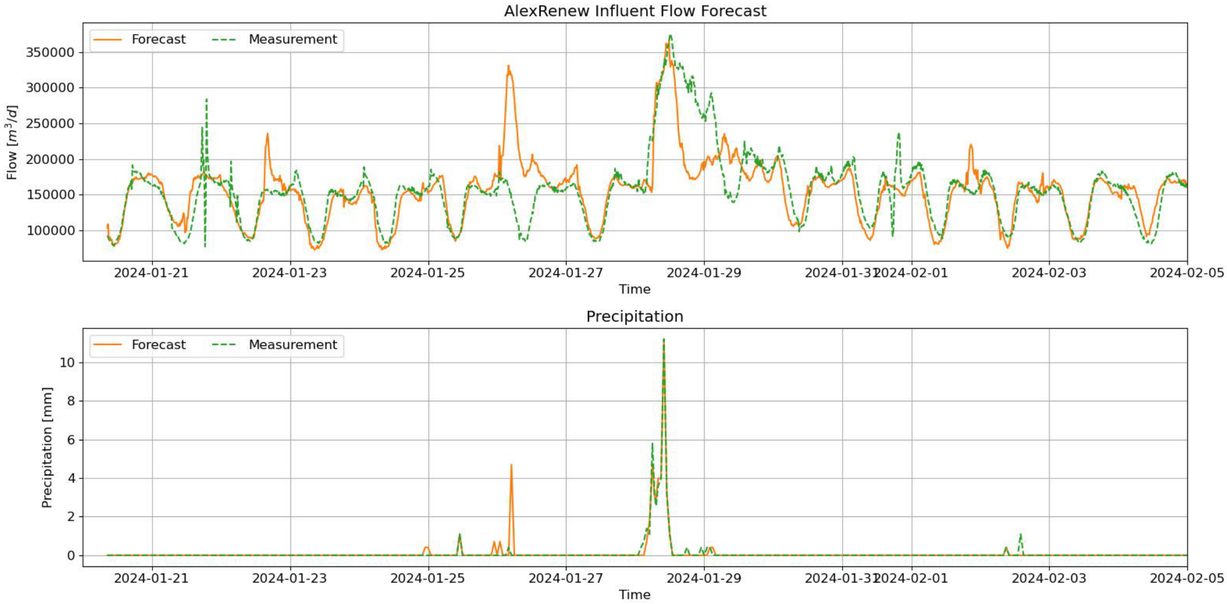

Figure 19: Influent flow forecast vs. measured flow for AlexRenew (top) and forecasted vs. measured precipitation (bottom).

As highlighted in on Figure 20 (zoomed in period highlighted in grey on Figure 18), the accuracy of the primary influent ammonia soft sensor concentration deteriorated slightly between January 26th and 30th. This can be linked to inaccuracies of the flow forecast model. However, comparing the resulting TKN load (Figure 21), the measured and simulated values consistently align with each other confirming the robustness of the applied methods for back-calculating influent variables. The impact on the mechanistic model is therefore limited to inaccurate hydraulics, which in this case is neglectable.

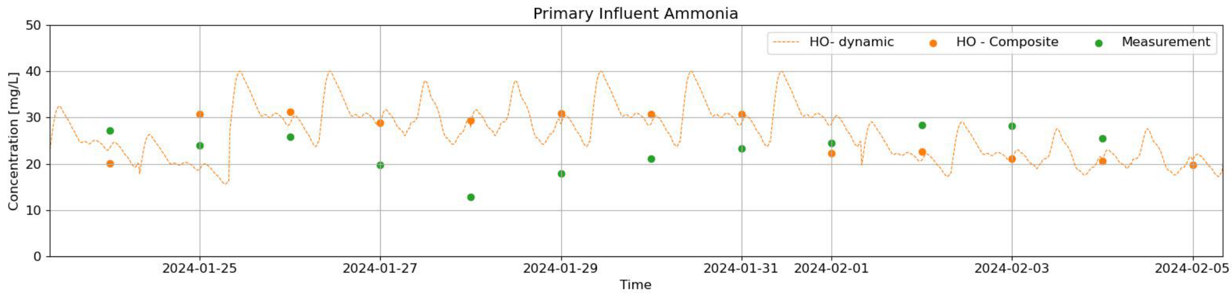

Figure 20: Comparison of measured (green) and simulated (orange) primary influent ammonia at AlexRenew WRRF. Zoomed-in period from January 24th to February 5th.

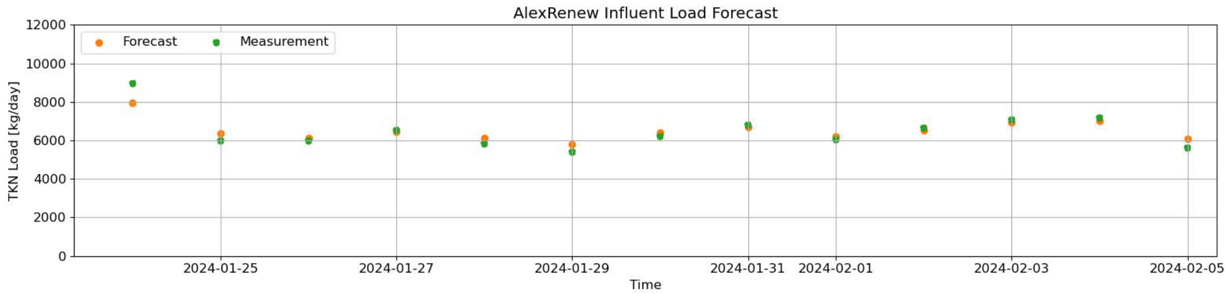

Figure 21: TKN Load Comparison of measured (green) and simulated (orange) primary influent at AlexRenew WRRF. Zoomed-in period from January 24th to February 5th.

Figure 22 is providing results for the control of the influent and effluent of the NMF with the goal to equalize the TKN load to the bioreactors. It should be noted that a requirement from the plant staff was that the NMF tank should be emptied once a day, which adds a constraint to the optimizer. The top plot shows the TKN load to the bioreactors under the current operation and if the Hybrid Optimizer recommendations are followed. The Hybrid Optimizer recommendations result in a significant lower standard deviation of 1,061 kg N/d, whereas the current operation results in a standard deviation of 1,350 kg N/d. The middle plot shows the flow to the bioreactors with the three flow setpoint recommendations per day. The highest setpoint is to empty the tank. The bottom plot shows the liquid volume of the NMF tank. It can be noted that the NMF tank could not be completely emptied during the given time frame. This has been caused by manual plant operation.

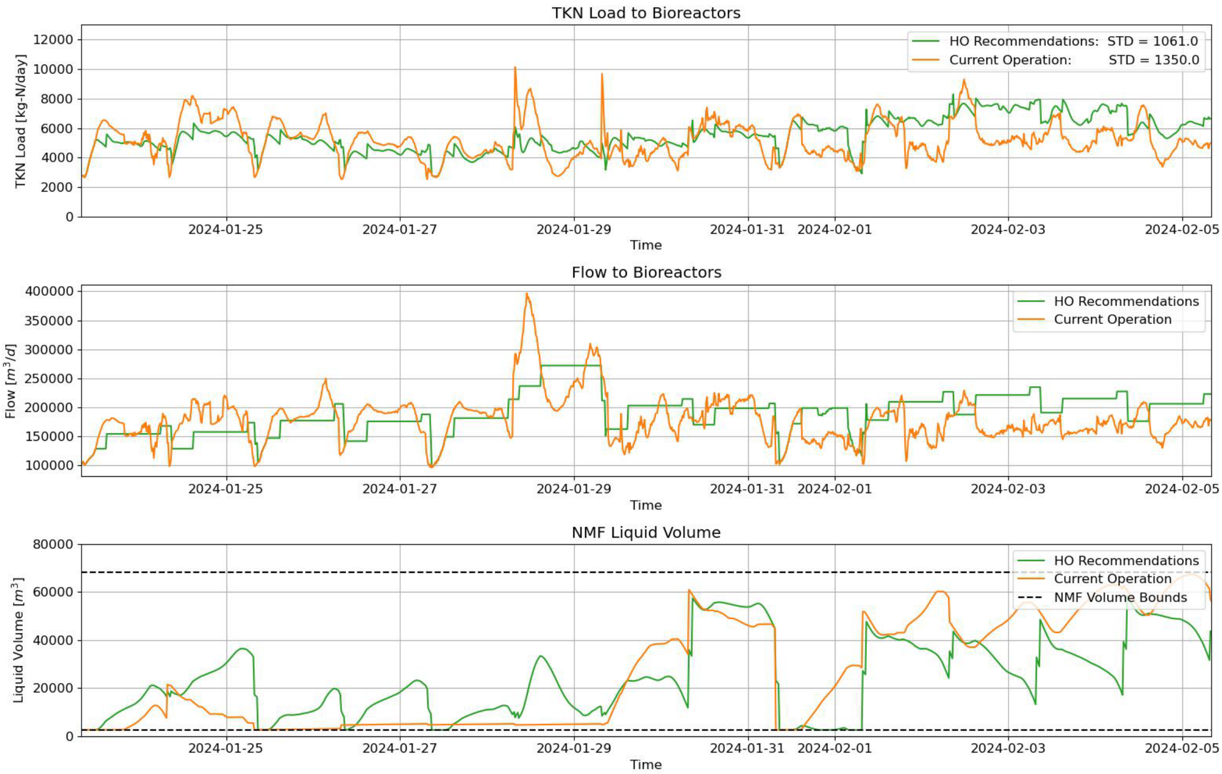

Figure 22: Performance of the Hybrid Optimizer controlling the AlexRenew Nutrient Management Facility (NMF) to equalize the TKN load to the bioreactors. Top Plot: TKN load to the bioreactor, Middle Plot: Recommended vs. actual flow to the bioreactors, and Bottom Plot: NMF tank volume.

The dashboards for AlexRenew are currently under development with a focus on providing insights to developers (Figure 23). The final dashboard setup will include the same structure for Operator, Analyst, and Engineer.

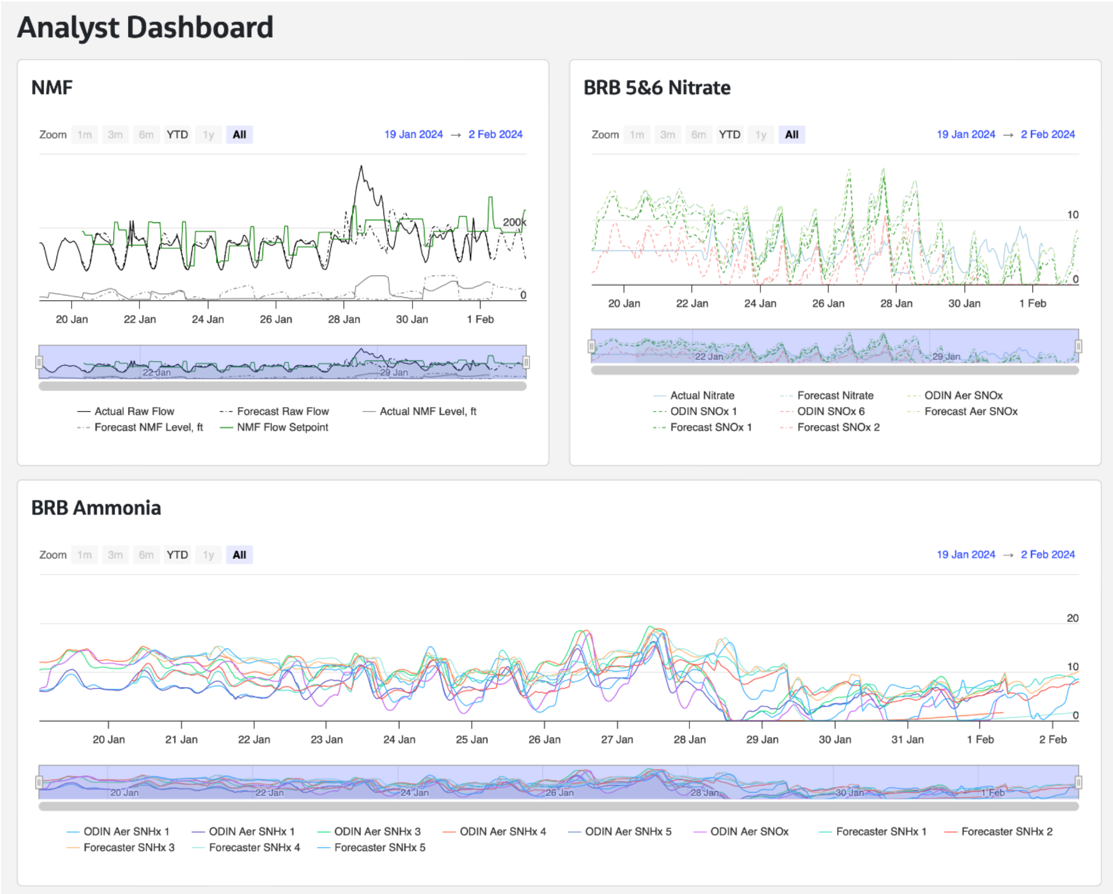

Figure 23: AlexRenew Analyst Dashboard Example.

Discussion and Lessons Learned

AlexRenew is a good example how a digital twin can be used to optimize facility operation. Without the soft sensor for the TKN load to the bioreactors, the only option would be to focus on pure flow control. Important for building trust in the Hybrid Optimizer models and recommendations was to regularly meet with plant staff show newest results and discuss next steps. The combination of digital tools and robust process engineering support was essential to move the project forward.

The flow forecast model was constantly improved during development. Important steps were the inclusion of river water levels to better integrate the effects of run-off and infiltration into the sewer system. The importance of data cleaning and filtering of measured data before use in a forecast model became obvious in a detailed analysis of the measured data and its impact on the forecast model.

Like the Agua Nueva WRF pilot, the Hybrid Optimizer at AlexRenew has led to several other insights to the facility. This has led to its use to help better understand DO control needs and effluent quality variations.

CONCLUSIONS

A novel concept combining the best of machine learning and mechanistic modelling for live digital twins in operational optimizations has been developed within The Water Research Foundation project 5121, Development of Innovative Predictive Control Strategies for Nutrient Removal. This paper provides results from three of the four case studies demonstrating the applicability and the benefits of the Hybrid Optimizer and a focus on usability by front-line operations staff. An interesting result of all three pilots was the benefits of the additional information provided by the soft sensor/mechanistic model assisting with other operational aspects of the facilities. The benefits can be summarized in four categories:

- ***Trustworthy, clean, and complete data sets*** due to a standardized data cleaning pipeline, and the combination of data-driven and mechanistic models. The mass-balancing of the mechanistic models adds a reliability layer to the results.
- ***Reduced Instrumentation Costs*** and Improved quality control with Soft Sensors.
- ***Front-line operations staff support*** though automation of data interpretation and ***Plant-wide insights*** based on all variables of interest.
- ***Integrated Data analytics tools*** provide unbiased evaluations of data adding to the process knowledge captured by the mechanistic models.
- ***Automated Cost and Reliability Optimization*** based on a comprehensive whole plant analysis.

REFERENCES

Belia, E., Neumann, M.B., Benedetti, L., Johnson, B., Murthy, S., Weijers, S., Vanrolleghem, P.A. (2021). Uncertainty in Wastewater Treatment Design and Operation, Uncertainty in Wastewater Treatment Design and Operation. IWA Publishing. doi.org/10.2166/9781780401034

Johnson, B.R., Yang, C., Lesnik, K., Registe, J., Johnson, T., Menniti, A. and Kenyon, J. (2023). Full-Scale Soft-Sensor Implementations Enable WRRF Hybrid Digital-Twins. Proceedings Watermatex 2023.

Menniti, A., Johnson, B.R. and Lesnik, K. (2023). Rise of the Machines: Applications of Machine Learning in Full-Scale Nutrient Management. Designing for Humans. Proceedings WEFTEC 2023.

Milton, J.S. and Arnold, J.C. (1995). Introduction to Probability and Statistics: Principles and applications for engineers and the computing sciences. Third Edition, McGraw-Hill, NY.

Registe, J., Lesnik, K., Johnson, B., Yang, C. and Pienta, D. (2023). Development of Innovative Predictive Hybrid Controller for Nutrient Removal. Proceedings WEFTEC 2023.

Regmi, P., Stewart, H., Amerlinck, Y., Arnell, M., García, P.J., Johnson, B., Maere, T., Miletić, I., Miller, M., Rieger, L., Samstag, R., Santoro, D., Schraa, O., Snowling, S., Takács, I., Torfs, E., van Loosdrecht, M.C.M., Vanrolleghem, P.A., Villez, K., Volcke, E.I.P., Weijers, S., Grau, P., Jimenez, J., Rosso, D. (2019). The future of WRRF modelling – outlook and challenges. Water Sci. Technol. 79(3–14). doi.org/10.2166/wst.2018.498

Rieger, L., Gillot, S., Langergraber, G., Ohtsuki, T., Shaw, A., Takács, I., Winkler, S. (2012). Guidelines for using Activated Sludge Models. IWA Scientific and Technical Report No. 22.

Schneider, M.Y., Quaghebeur, W., Borzooei, S., Froemelt, A., Li, F., Saagi, R., Wade, M.J., Zhu, J.J., Torfs, E. (2022). Hybrid modelling of water resource recovery facilities: status and opportunities. Water Sci. Technol. 85, 2503–2524. doi.org/10.2166/wst.2022.115

Therrien, J.-D., Nicolaï, N., Vanrolleghem, P.A. (2020). A critical review of the data pipeline: how wastewater system operation flows from data to intelligence. Water Sci. Technol. 82, 2613–2634. doi.org/10.2166/wst.2020.393

Yang, C., Johnson, B.R., Registe, J., Johnson, T., Rahman, A., Kenyon, J. (2023). Soft sensing influent concentrations with airflow rates by digital twins: A real case study from Oldham WwTW. Proceedings WEFTEC 2023.