A cost-effective IoT strategy for remote deployment of soft sensors – a case study on implementing a soft sensor in a multistage MBBR plant

A. M. Nair, A. Hykkerud and H. Ratnaweera

ABSTRACT

Model-based soft sensors can enhance online monitoring in wastewater treatment processes. These soft sensor scripts are executed either locally on a programmable logic controller (PLC) or remotely on a system with data-access over the internet. This work presents a cost-effective, flexible, open source IoT solution for remote deployment of a soft sensing algorithm. The system uses low-priced hardware and open-source programming language to set up the communication and remote-access system. Advantages of the new IoT architecture are demonstrated through a case study for remote deployment of an Extended Kalman Filter (EKF) to estimate additional water quality parameters in a multistage moving bed biofilm reactor (MBBR) plant. The soft-sensor results are successfully validated against standardised laboratory measurements to prove their ability to provide real-time estimations.

Key words | digital water, IoT, MBBR, soft sensor

A. M. Nair (corresponding author)
A. Hykkerud
H. Ratnaweera
Faculty of Science and Technology,
Norwegian University of Life Sciences,
1432 Ås,
Norway
E-mail: *muralabh@nmbu.no*

INTRODUCTION

The rise of digital water, driven by a rapid increase in information technology, has created an impetus to develop smarter and more efficient monitoring tools in wastewater treatment plants. Soft sensors have gained importance as a viable alternative to expensive online sensors (Olsson *et al.* 2014; Wang *et al.* 2019). Soft sensors are computer codes, which use a process model together with available online measurements from treatment plants to estimate additional wastewater quality parameters (Haimi *et al.* 2015). These soft sensor scripts are executed in a programmable logic controller (PLC), which has access to real-time data from the online sensors. Soft-sensors can also be implemented in a remote device if the plant's supervisory control and data acquisition (SCADA) has remote monitoring capabilities.

Most commercial SCADA vendors use proprietary software for programming their PLCs. Therefore, implementing the soft sensor in a PLC requires the code to be written in its proprietary programming suite, which limits the possibility of running soft sensor scripts written in commonly used scientific/academic programming languages such as Matlab or Python in industrial PLCs. The alternative is to develop a communication layer interface between a client (with the soft sensor script) and a server that has access to data from the online sensors. Open Platform Communication (OPC) has recently emerged as a global standard for communication between various sensors, PLCs and SCADA providers (González *et al.* 2017). Most SCADA vendors provide OPC servers with remote access options. These remote services do, however, come at an additional cost.

This work discusses an alternative, cost-effective IoT strategy, which provides flexible and secure remote access to the PLC. The new IoT architecture is implemented at a multistage moving bed biofilm reactor (MBBR) pilot plant to access the data from the online sensors remotely. Access to real-time data from the pilot enables remote deployment of an Extended Kalman Filter (EKF) to estimate additional water quality parameters in the plant.

MATERIAL AND METHODS

Pilot plant schematic and monitoring station

Figure 1 presents the operational schematic and monitoring station of a multistage MBBR pilot plant for municipal wastewater treatment (Saltnes *et al.* 2017). The reactor has

doi: 10.2166/wst.2020.067

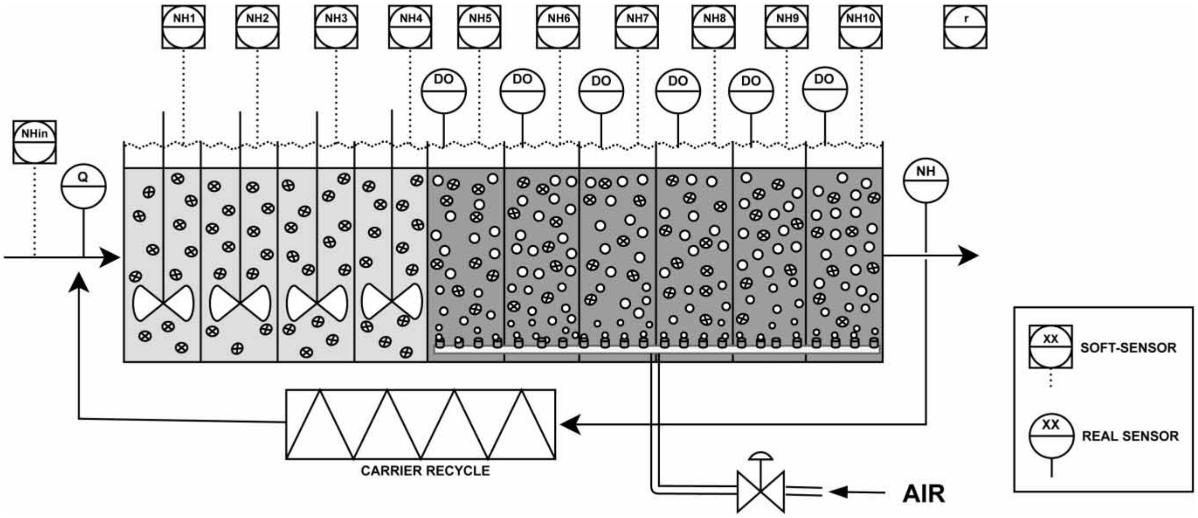

Figure 1 | Process flow diagram and sensor network in the pilot plant.

four anaerobic chambers, followed by six aerobic chambers. The plant is retrofitted with dissolved oxygen sensors from chamber 5 to 10 and an ammonia sensor in the effluent line. All online sensors are connected to a PLC provided by Beijer Electronics (https://www.beijerelectronics.no/). The current monitoring system is enhanced by deploying an array of soft sensors for monitoring additional water quality parameters. The aim of the soft-sensor array is to estimate the ammonia concentration in the influent raw wastewater as well as in each stage of the pilot plant. The soft sensor also provides estimates on additional parameters (such as the biomass activity of the nitrifiers) that cannot be measured directly using a physical sensor. Real-time data from soft sensors enable plant operators to adjust carrier recycle and aeration rate (manually or through control valves) to achieve optimal operation. The estimation of ammonia concentration in each aerobic chamber can enable the possibility of implementing ammonia-based aeration control (Rieger *et al.* 2013), which has proven to be a superior control strategy compared to a constant dissolved oxygen (DO) setpoint control (Vrečko *et al.* 2011).

Mathematical model for the multistage MBBR

Several mathematical models explaining biological nitrification are available in the literature. They vary from comprehensive models (Henze *et al.* 1987) used in design and optimization to simplified reduced-order models (Julien *et al.* 1999; Madyastha *et al.* 2011) used for monitoring and control. In this work, the nitrification kinetics for the biofilm process mentioned in Stare *et al.* (2006) are adapted for a multistage MBBR system. Equations (1)–(6) describe model equations in a discrete state-space form. Influent NH_4-N concentration ($S_{NH,in}$) and the nitrification rate (r) are estimated by augmenting them as additional state variables in the model.

$$x_{k+1} = x_k + T_s f(x, u) \tag{1}$$

$$x = [S_{NH,1}\ S_{NH,2}\ S_{NH,3}\ S_{NH,4}\ S_{NH,5}\ S_{NH,6}\ S_{NH,7}\ S_{NH,8}\ S_{NH,9}\ S_{NH,10}\ S_{NH,in}\ r] \tag{2}$$

$$u = [\,0 \quad 0 \quad 0 \quad 0 \quad S_{O,5} \quad S_{O,6} \quad S_{O,7} \quad S_{O,8} \quad S_{O,9} \quad S_{O,10}\,] \tag{3}$$

$$y = S_{NH,10} \tag{4}$$

$$f_i = \begin{cases} \tau^{-1}(S_{NH,in} - S_{NH,1}) - \rho_i, & i = 1 \\ \tau^{-1}\ (S_{NH,i-1} - S_{NH,i}) - \rho_i, & 1 \le i \le 10 \\ 0, & 11 \le i \le 12 \end{cases} \tag{5}$$

$$\rho_i = \begin{cases} 0, & i \le 4 \\ r * \dfrac{S_{NH,i}}{K_{NH} + S_{NH,i}} * \dfrac{1}{1 + e^{-K_1 S_{0,i} + K_2}}, & i \ge 5 \end{cases} \tag{6}$$

T_S is the sampling rate, k is the time subscript, and τ is the residence time in each chamber. The term ($S_{NH,i}$) represents the ammonia concentration and ($S_{o,i}$) represents the dissolved oxygen concentration in the ith chamber. The model-predicted value of the ammonia concentration in the effluent is denoted as y. The values of kinetic parameters K_1, K_2 and K_{NH} are taken from Stare *et al.* (2006).

Soft-sensor algorithm

The EKF is a widely used soft sensor algorithm for non-linear systems. Several examples for successful implementation of model-based soft sensors in biological wastewater treatment process using EKF can be found in the literature (Sotomayor *et al.* 2002; Busch *et al.* 2013; Zeng *et al.* 2016). The basic equations in an EKF are presented in Equations (7)–(13).

$$F_{\mathrm{k}} = I + T_{\mathrm{s}} \frac{\partial f}{\partial x}\bigg|_{x_{\mathrm{k}},\, u_{\mathrm{k}}} \tag{7}$$

$$x^{-}_{\mathrm{k}+1} = x_{\mathrm{k}} + T_{\mathrm{s}} f(x_{\mathrm{k}}, u_{\mathrm{k}}) \tag{8}$$

$$P^{-}_{\mathrm{k}+1} = F_{\mathrm{k}-1} P_{\mathrm{k}} F^{T}_{\mathrm{k}} + Q \tag{9}$$

$$H_{\mathrm{k}} = \frac{\partial h}{\partial x}\bigg|_{x^{-}_{\mathrm{k}+1}, u_{\mathrm{k}}} \tag{10}$$

$$K_{\mathrm{k}} = P^{-}_{\mathrm{k}+1} H^{T}_{\mathrm{k}} (H_{\mathrm{k}} P^{-}_{\mathrm{k}+1} H^{T}_{\mathrm{k}} + R)^{-1} \tag{11}$$

$$x_{\mathrm{k}+1} = x^{-}_{\mathrm{k}+1} + K_{\mathrm{k}}(z_{\mathrm{k}} - h(x^{-}_{\mathrm{k}+1})) \tag{12}$$

$$P_{\mathrm{k}+1} = (I - K_{\mathrm{k}} H_{\mathrm{k}})\, P^{-}_{\mathrm{k}+1} \tag{13}$$

In the equations above, x_k is the state vector and z_k is the measurement vector at a time instance k. The process and measurement noise covariance matrices are represented as Q and R respectively. I is the identity matrix, x^{-}_{k+1} is the *a priori* estimate of the state, P^{-}_{k+1} the covariance of *a priori* estimation error and P_k the covariance matrix of the *a posteriori* estimation error. The sequential order of executing the EKF equations is presented as a flowchart in Figure 3(b).

Before implementing the soft-sensor in the pilot plant, the mathematical model (Equations (1)–(6)) and the EKF (Equations (7)–(13)) were implemented on a simulator platform to assess the system observability and identify the minimum number of online sensors required to estimate the ammonia concentration in each chamber. The simulator-based testing strategy implemented in Nair *et al.* (2019) was

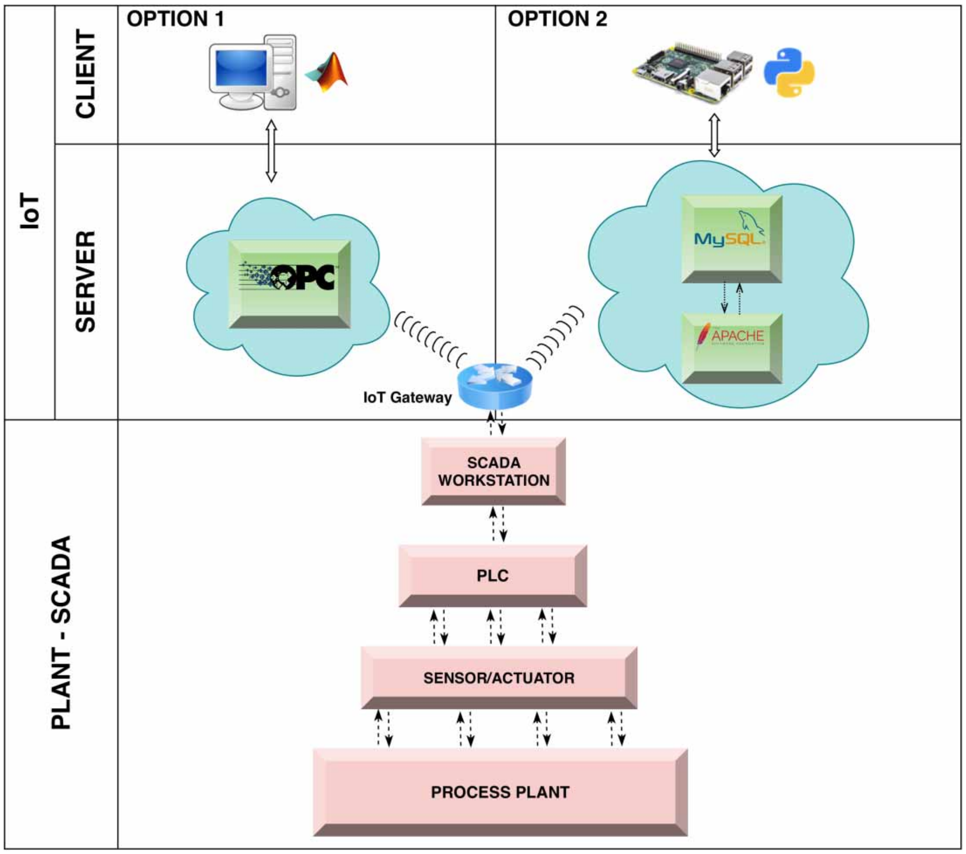

Figure 2 | IoT schematic option 1 (OPC) and option 2 (MySQL).

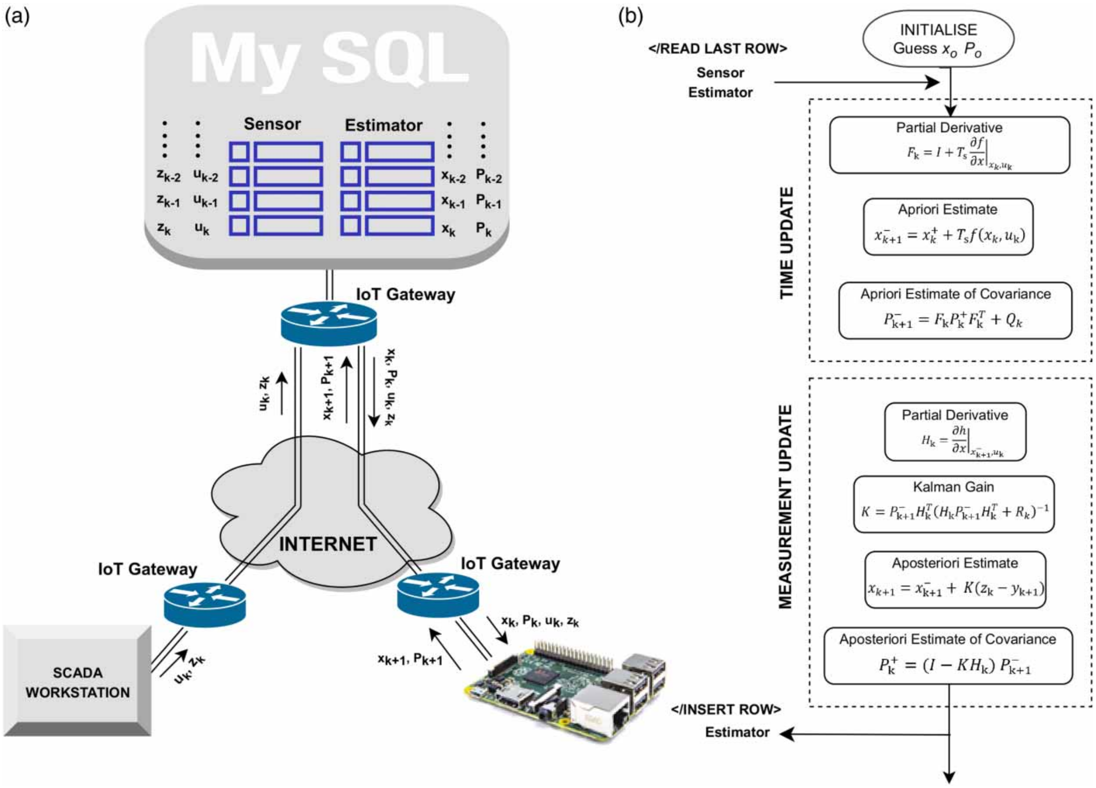

Figure 3 | (a) Data flow during remote deployment of EKF; (b) flowchart for EKF implementation.

used to assess the performance of EKF. The values of the tuning parameters of EKF; Q, R, P_0 x_0 are obtained using the tuning methods described in Haugen *et al.* (2014).

IoT schemes and soft sensor deployment

Figure 2 presents two possible schematics for remote deployment of the soft sensor code. Option 1 presents the default system provided by the SCADA vendor, which consists of an OPC server accessible through a secure VPN connection. The EKF is written as a script in Matlab and is executed remotely on a PC. The OPC Client toolbox in Matlab is used to receive real-time sensor data from the OPC servers. The second option includes an IoT gateway, which forms an interface between the PLC and a remote MySQL server, which is made accessible using Dynamic DNS (DDNS). The soft sensor algorithm is written as a Python script, which is then deployed on a single board computer – Raspberry Pi (https://www.raspberrypi.org/).

The network architecture for the remote deployment of the soft sensor is presented in Figure 3(a). The IoT gateway at the pilot plant collects data from online sensors and inserts them as a new row in a table created in the remote MySQL server. The client running the EKF code (in the Raspberry Pi) begins the cycle by reading the last row of the table. This table consists of the estimated states x_k, the covariance matrix of estimation error P_k, measurements from online sensor z_k and u_k. The EKF algorithm then calculates new values of the estimated state x_{k+1} and covariance matrix of estimation error P_{k+1}. The updated values are inserted as a new row in the MySQL table. The Python script for implementing an EKF on a non-linear state-space model, the communication code that enables read and write of data from the remote SQL database, and the associated configuration files are provided in the Supplementary Material. The code has a modular structure and can be easily adapted to a new system by anyone with basic knowledge of programming in Python.

RESULTS AND DISCUSSION

Soft-sensor validation

The estimation results for a period of 1 month (28th January 2019–12th March 2019) are presented in the following

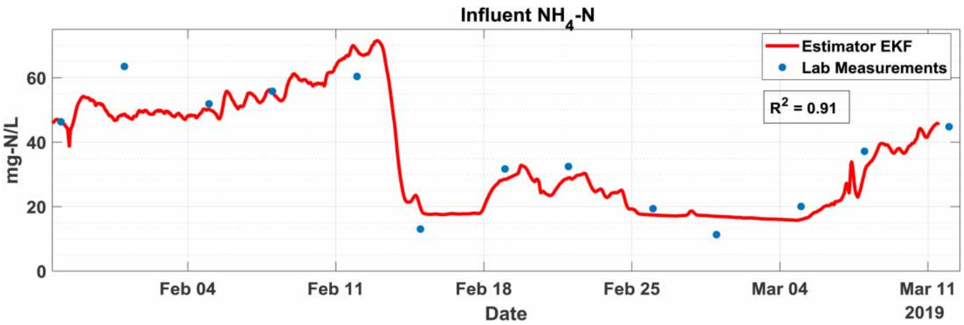

Figure 4 | Soft sensor results for a duration of 1 month. EKF vs. laboratory measurements.

figures. Figure 4 shows the comparison between the influent NH_4-N concentration estimated by soft sensor and the biweekly measurements obtained by standardized laboratory tests (Hach-Lange kits). The R^2 values presented in the graph show that the values predicted by the estimator are close to the ammonia measurements from the laboratory. This demonstrates the effectiveness of soft sensors in providing real-time values of influent NH_4-N concentration.

The estimation results for NH_4-N concentrations at various stages in the pilot plant are presented in Figure 5. The time span for ammonia concentration in the individual chamber is shortened to one week for better visualisation of data.

Figure 6 presents the value of the nitrification rate *r* estimated by the soft sensor. A simplified nitrification model, such as the one used in this work, is derived from the comprehensive activated sludge model by considering certain assumptions. One such assumption involves excluding the kinetics of biomass growth and death and considering a constant biomass concentration (Julien *et al.* 1999). However, a time-varying estimate of the nitrification rate in the pilot could be a result of variations in biomass activity. Therefore, the real-time estimation of the time-varying parameter *r* could provide significant insight on variations in the concentration of active nitrifiers in the pilot plant. A qualitative assessment of the trends (presented in Figure 6) indicates a positive correlation between the nitrification rate and the ammonia concentration of influent wastewater. The nitrification rate (*r*) showed a steady decline on February 13th, a few hours after the influent ammonia concentration dropped. The pilot plant showed a lower nitrification rate from February 15th until March 6th, when the influent ammonia concentration was low. The nitrification rate started to increase after March 6th, when the influent ammonia concentration started to rise. This observation corresponds to the usual behavior of nitrifiers present in a biological wastewater treatment plant. However, a detailed analysis of the estimated results and cross-validation against

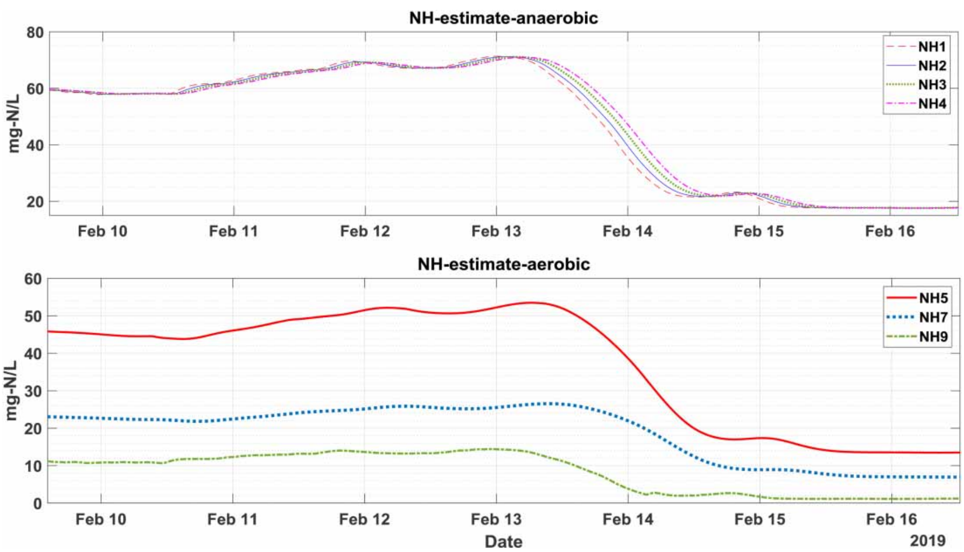

Figure 5 | Soft sensing results for NH_4-N concentration in anaerobic chamber (top) and aerobic chamber (bottom).

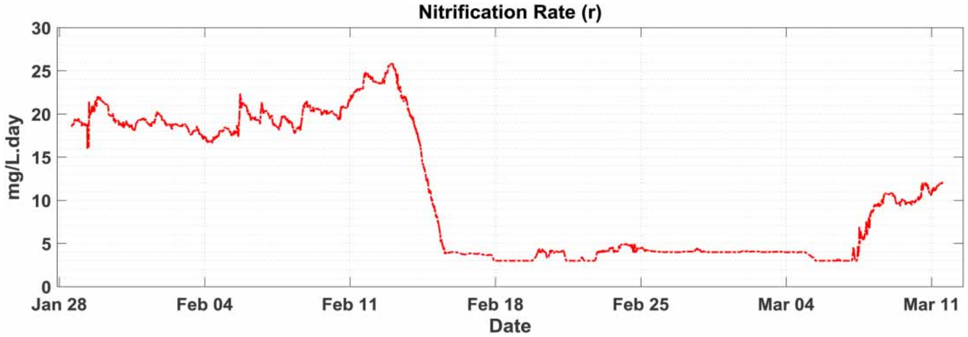

Figure 6 | Soft sensing results for nitrification rate.

laboratory measurements should be performed to substantiate the claim and fine-tune the estimator for soft sensing the concentration of autotrophs in the pilot plant.

Cost comparison

A cost comparison between the default option provided by the SCADA vendor (Option 1) and two different versions (Option 2 and Option 3) of the new IoT system is presented in Table 1. Table 1 also presents the cost incurred while using a physical ammonia sensor (ion selective electrode) for measuring ammonia concentration in all six aerobic chambers. It can be inferred from the overall cost that using a soft-sensor can result in significant savings.

Among the options for soft-sensors, the default option provided by the SCADA provider has an average cost of about 1,580 € compared to the new system, which costs about 155 € (with SSL certificates). There is the possibility of running the setup without SSL, which would bring the cost down to 35 €, but this is not recommended due to security concerns (Buchanan *et al.* 2017). Recently there are providers that supply free SSL certificates (Aertsen *et al.* 2017), but with a slightly lower security level than the commercial options due to limited support. This could be a valid option for security if the system is not critical and other operational security measures like limited access control levels and network monitoring are taken into consideration. For more critical water infrastructure, additional layers of security can be added to the communication network by providing a VPN or an intermediate API link. The open source VPN software OpenVPN is relatively easy to implement, but API links require significant design considerations and can become inflexible. The new system discussed in this work provides a cost-effective option where the only investment is in the hardware (Raspberry Pi) and the cost associated with the purchase of commercial services for establishing a secure data connection.

CONCLUSIONS

This work illustrates a functioning example of a cost-effective and flexible alternative for non-intrusive remote

Table 1 | Cost comparison between option 1 (default) and options 2 and 3 (alternative IoT strategy) option 4 (physical sensor)

	Option 1		**Option 2**		**Option 3**		**Option 4**	
	Service	**Cost (€)**	**Service**	**Cost (€)**	**Service**	**Cost (€)**	**Service**	**Cost (€)**
Server	OPC (Codesys)	160	MySQL	–	MySQL	–	–	
Remote access	VPN (Insys)	560	DDNS	–	DDNS	–	–	
Security/encryption	Inbuilt	–	SSL key	120	–	–		
Client hardware	PC	200–1,000	Raspberry Pi	35	Raspberry Pi	35	–	
Client software	Matlab	220	Python	–	Python	–	–	
Client software	Matlab	220	Python	–	Python	–	–	
NH_4-N sensor (ISE)	1 unit	1,500	1 Unit	1,500	1 Unit	1,500	6 Units	1,500
IoT cost		980–1,940		155		35		0
Overall cost		2,480–3,440		1,655		1,535		9,000

monitoring in a wastewater treatment plant. The soft-sensor array deployed in the pilot plant proves the ability to bolster the monitoring system in a treatment process. The ability to monitor additional wastewater treatment parameters that are otherwise difficult to monitor via conventional online sensors would enable the implementation of multi-parameter based optimal control strategies. The case study presented in this work also illustrates the possibility of integrating scripts written in commonly used scientific programming languages such as Matlab or Python into any commercial SCADA. The cost comparison between the new IoT system and the options available in the market today shows that remote deployment of soft sensors can be achieved at much lower costs. This novel cost-effective option also opens up new possibilities to develop more comprehensive soft-sensing algorithms for estimating water quality parameters, which are otherwise difficult to monitor by a physical sensor.

ACKNOWLEDGEMENTS

We acknowledge the financial support provided by Norwegian Research Council (RECOVER project) and the technical support from the research team at Norwegian University of Science and Technology (NTNU).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this paper is available online at https://dx.doi.org/10.2166/wst.2020.067.

REFERENCES

Aertsen, M., Korczynski, K., Moura, C. M. G., Tajalizadehkhoob, S. & Berg, J. 2017 No domain left behind: is Let's Encrypt democratizing encryption? *ANRW* **17**, 48–54.

Buchanan, W. J., Helme, S. & Woodward, A. 2017 Analysis of the adoption of security headers in HTTP. *IET Information Security* **12** (2), 118–126.

Busch, J., Elixmann, D., Kühl, P., Gerkens, C., Schlöder, J. P., Bock, H. G. & Marquardt, W. 2013 State estimation for large-scale wastewater treatment plants. *Water Research* **47** (13), 4774–4787.

González, I., Calderón, A., Barragán, A. & Andújar, J. 2017 Integration of sensors, controllers and instruments using a novel OPC architecture. *Sensors* **17** (7), 1512.

Haimi, H., Corona, F., Mulas, M., Sundell, L., Heinonen, M. & Vahala, R. 2015 Shall we use hardware sensor measurements or soft-sensor estimates? Case study in a full-scale WWTP. *Environmental Modelling and Software* **72**, 215–229.

Haugen, F., Bakke, R. & Bernt, L. 2014 State estimation and model-based control of a pilot anaerobic digestion reactor. *Journal of Control Science and Engineering* **2014** (3), 1977–1996.

Henze, M., Grady, C. P. L., Gujer, W., Marais, G. v. R. & Matsuo, T. 1987 *Activated Sludge Model No.1*. IAWPRC Scientific and Technical Reports No. 1. IWA Publications, London.

Julien, S., Lessard, P. & Barbary, J. P. 1999 A reduced order model for control of a single reactor activated sludge process. *Mathematical and Computer Modelling of Dynamical Systems* **5** (4), 337–350.

Madyastha, V. K., Prasad, V. & Mahendraker, V. 2011 Reduced order model monitoring and control of a membrane bioreactor system via delayed measurements. *Water Science & Technology* **64** (8), 1675–1684.

Nair, A. M., Fanta, A., Haugen, F. A. & Ratnaweera, H. 2019 Implementing an Extended Kalman Filter for estimating nutrient composition in a sequential batch MBBR pilot plant. *Water Science and Technology* **80** (2), 317–328.

Olsson, G., Carlsson, B., Comas, J., Copp, J., Gernaey, K. V., Ingildsen, P., Jeppsson, U., Kim, C., Rieger, L., Rodriguez-Roda, I., Steyer, J. P., Takács, I., Vanrolleghem, P. A., Vargas, A., Yuan, Z. & Åmand, L. 2014 Instrumentation, control and automation in wastewater – from London 1973 to Narbonne 2013. *Water Science and Technology* **69** (7), 1373–1385.

Rieger, L., Jones, R. M., Dold, P. L. & Bott, C. B. 2013 Ammonia-based feedforward and feedback aeration control in activated sludge processes. *Water Environment Research* **86** (1), 63–73.

Saltnes, T., Sørensen, G. & Eikås, S. 2017 Biological nutrient removal in a continuous biofilm process. *Water Practice and Technology* **12** (4), 797–805.

Sotomayor, O. A., Park, S. W. & Garcia, C. 2002 Software sensor for on-line estimation of the microbial activity in activated sludge systems. *ISA Transactions* **41** (2), 127–143.

Stare, A., Hvala, N. & Vrečko, D. 2006 Modeling, identification, and validation of models for predictive ammonia control in a wastewater treatment plant–a case study. *ISA Transactions* **45** (2), 159–174.

Vrečko, D., Hvala, N. & Stražar, M. 2011 The application of model predictive control of ammonia nitrogen in an activated sludge process. *Water Science and Technology* **64** (5), 1115–1121.

Wang, X., Kvaal, K. & Ratnaweera, H. 2019 Explicit and interpretable nonlinear soft sensor models for influent surveillance at a full-scale wastewater treatment plant. *Journal of Process Control* **77**, 1–6.

Zeng, J., Liu, J., Zou, T. & Yuan, D. 2016 Distributed extended Kalman filtering for wastewater treatment processes. *Industrial & Engineering Chemistry Research* **55** (28), 7720–7729.

First received 31 October 2019; accepted in revised form 6 February 2020. Available online 18 February 2020