Nitrate estimation in the denitrifying post-filtration unit of a municipal wastewater treatment plant: the Viikinmäki case

M. Mulas, F. Corona, H. Haimi, L. Sundell, M. Heinonen and R. Vahala

ABSTRACT

In this work we present and discuss the design of an array of soft-sensors to estimate the nitrate concentration in the denitrifying post-filtration unit at the Viikinmäki wastewater treatment plant in Helsinki (Finland). The developed sensors aim at supporting the existing hardware analyzers by providing a reliable back-up system in case of malfunction of the instruments. In the attempt to design easy to implement and interpretable sensors, computationally light linear models have been considered. However, due to the intrinsic nonlinearity of the process, also nonlinear but still computationally affordable models have been considered for comparison. The experimental results demonstrate the potential of the developed soft-sensors and the possibility for an on-line implementation in the plant's control system as alternative monitoring devices.

Key words | post-denitrification, process monitoring and supervision, soft-sensors, wastewater treatment

M. Mulas (corresponding author)
H. Haimi
R. Vahala
Aalto University,
Dept. of Civil and Environmental Engineering,
P.O. Box 15200, FI-00076 Aalto,
Finland
E-mail: *michela.mulas@aalto.fi*

F. Corona
Aalto University,
Dept. of Information and Computer Science,
P.O. Box 15400, FI-00076 Aalto,
Finland

L. Sundell
M. Heinonen
HSY Helsinki Region Environmental Services Authority,
P.O. Box 100, FI-00066 HSY,
Finland

INTRODUCTION

The observance of more and more stringent environmental regulations, the need for sustainable reuse of water and the optimization of costs represent unceasing challenges for the operation and management of municipal wastewater treatment plants (WWTPs). Accordingly, plants are evolving toward efficient and safe operations with high-quality effluents, while optimizing operating and management costs. A prime requirement for achieving these objectives relies on real-time automation technologies allowing us to efficiently monitor and supervise process units, to implement advanced control strategies and, on a higher level, to achieve plant-wide optimization (Olsson *et al.* 2005).

Nowadays, due to the advances in measuring technology, an increased amount of on-line and off-line measurements from hardware sensors, such as, field and laboratory analyzers, is routinely used to monitor WWTP processes. When on-line instrumentation is available, however, one of the major problems is that it cannot always be relied upon, mainly because of the harsh environment. Partial failure in the sensors (e.g. due to measurement off-set, drift or precision degradation) may result in a false perception of the performances of the plant's units and erroneous control actions. Consequently, the need for systems that supply the plant's operators and engineers with decision-support information on the instrumentation has also increased. In such a context, software sensors, or soft-sensors, are a complementary and inexpensive key technology when hardware sensors are unreliable or not available (Dochain & Vanrolleghem 2001; Kadlec *et al.* 2009).

Extensively used in the process and power industry, soft-sensors allow estimation in real-time of some hard-to-measure but important process variables and quality indicators (primary or output variables), starting from some other easy-to-measure ones (secondary or input variables). Conventionally, data-derived soft-sensors are designed as input–output models combining regression and fault-detection methods borrowed from chemometrics and multivariate statistical process control (for reference, see Rosen 2001; Vanrolleghem & Lee 2003; Aguado & Rosen 2008; Lee *et al.* 2008; Yoo *et al.* 2008; Baggiani & Marsili-Libelli 2009; Gibert *et al.* 2010 and Avella *et al.* 2011). Especially in the last decades, the increased availability of process data also in wastewater treatment facilities has thus motivated

doi: 10.2166/wst.2012.038

the effort toward the definition of systematic approaches for the development of data-derived soft-sensors.

In this paper, we illustrate the development of an array of soft-sensors for real-time estimation of nitrate concentrations in the post-denitrification unit of the Viikinmäki WWTP in Helsinki, Finland. The harsh environment where the hardware sensors operate exposes their analysis to potential malfunctioning. The availability of accurate nitrate concentration measurements is, on the other hand, fundamental from both an environmental and an economical point of view. Thus a back-up software system capable of complementing the existing instrumentation is needed to enforce optimality in the operation of the unit.

PROCESS DESCRIPTION

The Viikinmäki WWTP (800,000 Population Equivalent) is the largest WWTP in Finland; it is built inside a bed rock and treats an average influent flow rate of 250,000 m^3/d with peaks of 800,000 m^3/d. The wastewater treatment line of the plant consists of bar screening, grit removal, pre-aeration, primary sedimentation, activated sludge process (divided into eight treatment lines), secondary sedimentation and a biological post-filtration (post-denitrification). The sludge treatment is achieved with four mesophilic digesters and subsequent dewatering systems. The biogas from the sludge digestion is utilized for electricity and heat production.

Due to strict treatment requirements on nitrogen compounds, the plant has been updated twice since its start-up in 1994; in particular, since 2004 with the introduction of tertiary filters, a total nitrogen removal level of approximately 90% of yearly average has been achieved. The post-denitrification process receives wastewater from the secondary sedimentation unit and nitrate removal is achieved by means of ten Biostyr™ filters arranged in parallel, Figure 1(a). The influent wastewater is equally distributed to ten filter cells and, before each cell, the incoming flow is split in to two separate streams in which methanol diluted in water, with a constant 10% dilution, is added independently to each line to favor the removal, Figure 1(b). The diluted methanol flow rate is measured in each line (*Fi-QM-1/2*) and the overall flow rate to each filter is manipulated through a feedback loop controlling the nitrate-nitrogen concentration (*Fi–NO*$_3$) in the cell. The nitrate-nitrogen concentration in the filter is measured in real-time by means of an optical instrument (Nitratax™ plus). Inside the cell, wastewater flows upwards through a floating support media, filling approximately half of the bed, where biomass is attached. Periodic backwashes are needed at intervals ranging from 12 to 72 h according to the head-loss (*Fi–HL*) due to the biomass attached to the support media that tends to clog the cell. The cells are usually backwashed, one at a time, using the effluent wastewater (*Fi-QWW*) with a counter-current airflow (*Fi-QWA*). The average retention time of the filtering process is 25 minutes. After filtration, the wastewater treated in the filter is discharged into the effluent channel where streams from Filter 10 to 1 are collected.

The quality of the wastewater entering the filters is monitored on-line, before division to each filter cell, in terms of dissolved oxygen DO (*I-O*$_2$), suspended solids SS (*I-SS-1/2*), nitrate-nitrogen NO_3-N (*I-NO*$_3$*-1/2*), phosphate PO_4-P (*I-OP*) and total phosphorus (*I-TP*). As for the effluent treated wastewater, temperature (*E-T*), total organic carbon (*E-TOC*), nitrate-nitrogen (*E-NO*$_3$), phosphate (*E-OP*) and total phosphorus (*E-TP*) are measured on-line in the discharge channel. In addition, laboratory analyses of daily averaged samples are performed twice per week, for nutrient (nitrogen and phosphorus compounds), suspended solids and organic compounds in the influent and effluent wastewater. The number of process variables measured in real-time is 142: 7 for the influent, 5 for the effluent and 13×10 for the filters, as summarized in Table 1.

Although the facility does not automatically register this information, according to the plant's management, the on-line measurements of nitrate in the filter cells are often disturbed by the frequent backwashes and, thus, they are prone to potential unreliability. One principal problem is the vicinity of the analyzers to the effluent channel; in fact, the effect of a backwash in one upstream filter (say, Filter 7) propagates to the filters downstream (that is, Filters 6 to 1). Moreover, all the instruments are exposed to an aggressive environment. On the other hand, a reliable analysis of these important variables is mandatory because of the economical and environmental implications due to an incorrect dosing of methanol. For such reasons, we investigate the possibility to develop an array of soft-sensors (one for each filter) that estimate in real-time the nitrate-nitrogen concentration in the cells. The design of simple, intelligible and computationally light regression models is a priority for direct implementation in the plant's control system (Kano & Ogawa 2010).

SOFT-SENSOR DEVELOPMENT

To develop the sensors, a set of process measurements from the plant's data acquisition system has been collected.

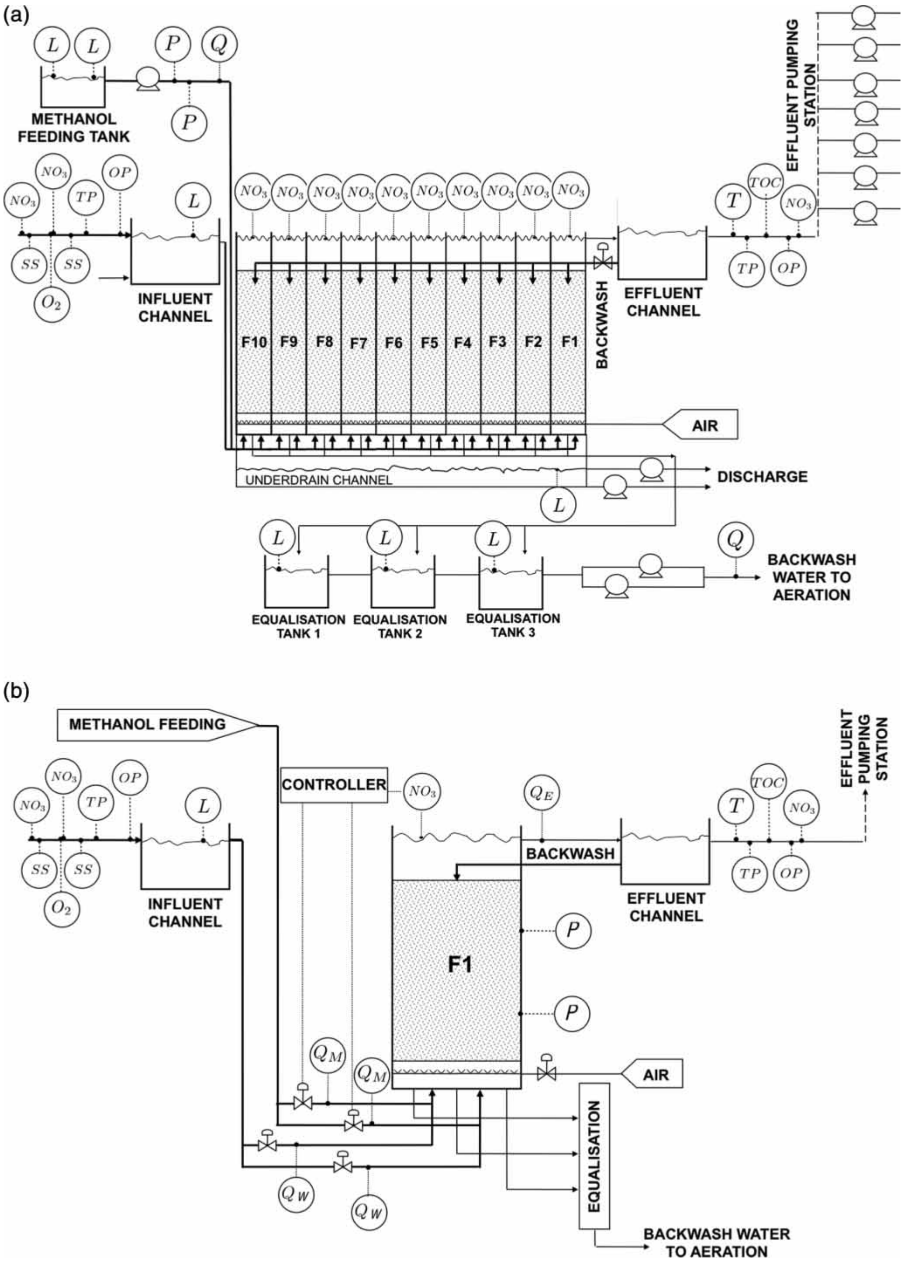

Figure 1 | A schematic representation of the post-denitrification unit of the Viikinmäki WWTP (a) with highlight on one filter (b).

The data correspond to three years of continuous operations (Jan 1, 2008–Dec 31, 2010), recorded as hourly averages. Measurements from the first year have been used for the exploratory data analysis (observation and variable selection) and for calibrating the regression models, whereas measurements from the second and third year are then used as an independent set for testing the models' performance. In the following, the main stages of the soft-sensors' design are discussed, starting from the selection of reliable output samples to learn the regression models and the selection of a subset of relevant input variables and ending with the calibration of the models.

Sample selection

Given the poor quality of the measurements of nitrate-nitrogen concentration in the filters, the first stage consisted of performing a selection of the samples to be used to calibrate the regression models. The task has been approached

Table 1 | The complete list of the process variables of the post-denitrification unit of the Viikinmäki WWTP

I-NO_3-1	Influent NO_3-N (sensor 1)	mg/L
I-NO_3-2	Influent NO_3-N (sensor 2)	mg/L
I-SS-1	Influent SS (sensor 1)	mg/L
I-SS-2	Influent SS (sensor 2)	mg/L
I-O_2	Influent DO	mg/L
I-OP	Influent PO_4-P	mg/L
I-TP	Influent total P	mg/L
Fi-QWW	*i-th* Filter backwash water	m^3/s
Fi-QWA	*i-th* Filter backwash air	m^3/s
Fi-QW-1	*i-th* Filter wastewater (line 1)	m^3/s
Fi-QW-2	*i-th* Filter wastewater (line 2)	m^3/s
Fi-QM-1	*i-th* Filter methanol (line 1)	m^3/h
Fi-QM-2	*i-th* Filter methanol (line 2)	m^3/h
Fi-P-1	*i-th* Filter bottom pressure	kPa
Fi-P-2	*i-th* Filter top pressure	kPa
Fi-NO_3	*i-th* Filter NO_3-N	mg/L
Fi-HL	*i-th* Filter head-loss	m
Fi-CR	*i-th* Filter clogging rate	%
Fi-HRU	*i-th* Filter hours in use	–
Fi-ITW	*i-th* Filter intermediate time of backwash	–
E-NO_3	Effluent NO_3-N	mg/L
E-TOC	Effluent TOC	mg/L
E-OP	Effluent PO_4-P	mg/L
E-TP	Effluent total P	mg/L
E-T	Effluent Temperature	°C

starting from the consideration that, in normal operating conditions, the filters should operate in ideally identical ways. This implies that the nitrate-nitrogen measurements in different filters should be very similar. As a consequence, deviations from prime operational modes are to be associated with instrumental anomaly and these erroneous measurements should be discarded.

Such deviations can be readily detected by analyzing the data matrix $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_i, \ldots, \mathbf{y}_{10}]$, where $\mathbf{y}_i$ denotes a vector of nitrate-nitrogen concentrations *Fi-NO_3* in the *i*-th filter. For the sake of clarity, note that $\mathbf{Y}$ contains only the K samples that correspond to normal operations; that is, when the filter is not-backwashing and, for each filter, when none of the filters upstream is backwashing. Given the ideal collinearity between the $\mathbf{y}_i$ variables, $\mathbf{Y}$ is expected to be characterized by a very high (ideally, one) and flat correlation matrix and by very low dimensionality (ideally, the intrinsic dimensionality is one). This is confirmed by the analysis of the correlation matrix for Filters 4 to 8, whereas this structure is still strong but slightly deteriorated for Filters 2, 3 and 9. Conversely, Filters 1 and 10 are characterized by a behavior that is highly different from the other cells, already suggesting a potential malfunctioning of the corresponding instruments. As for the intrinsic dimensionality, deviations can be detected by analyzing the variance-covariance structure of $\mathbf{Y}$ with a Principal Component Analysis model, PCA (Jolliffe 2002). By PCA, the $K \times I$ matrix $\mathbf{Y}$ is factorized into a $K \times S$ *score matrix* $\mathbf{T}$ and an $I \times S$ *loading matrix* $\mathbf{P}$ by eigenvalue decomposition of the covariance matrix of $\mathbf{Y}$, to obtain $\mathbf{Y} = \mathbf{TP}' + \mathbf{E}$. $\mathbf{E}$ is the $K \times I$ *residual matrix*.

In our case, small residuals are already expected for small values of S and this is particularly true for Filters 4 to 8. For such a reason, to select a number S of *principal components* (PCs) for a PCA model that appropriately reconstructs only the main modes of variation in the filters, but discards instrumental deviations, the reconstruction error for each filter is analyzed for a varying number of components using the *Root Mean Squared Error* (*$RMSE_i$*) for the *i*-th filter:

$$\mathrm{RMSE}_i = \left(\frac{1}{K} \sum_{k=1}^{K} (y_i(k) - \hat{y}_i(k))^2 \right)^{1/2} \quad \text{with } i = 1, \ldots 10 \quad (1)$$

where $y_i(k)$ and $\hat{y}_i(k)$ denote the nitrate-nitrogen measurement in the *i*-th filter at time k and its reconstruction from the PCA model, respectively. The results with only $S = 2$ components showed that the signals corresponding to Filters 4 to 8 are already well recovered, thus confirming the intrinsically low dimensionality of the process. As for Filters 1 and 10, a larger number of components would be required to reconstruct the residual variability. For such reasons, a bi-dimensional PCA model capable of explaining 60% of the total data variation has been considered. On such a model, samples can be selected based on the residuals of the PCA model. To analyze the residuals, two commonly used figures of fit have been considered: the *Squared Prediction Error*, *SPE*, or *Q-statistic*, and the *Hotelling's T^2 statistic*. The *SPE* is used to measure the distance of an observation from its reconstruction on the PC subspace and T^2 measures the (normalized) distance of the projected observation $\mathbf{t}(k)$ from the origin of the PC subspace. That is,

$$SPE(k) = \sum_{i=1}^{I} (y_i(k) - \hat{y}_i(k))^2 \quad \text{and} \quad T^2(k) = \mathbf{t}(k)\Lambda^{-1}\mathbf{t}'(k) \quad (2)$$

where $\mathbf{\Lambda}^{-1}$ denotes a diagonal matrix with the inverse of the eigenvalues for the retained PCs. The two statistics are combined into a single index, the *J-statistic* (Raich & Cinar 1996), which quantifies how much an observation globally disagrees, or *uncorrelates*, with the PCA model:

$$J(k) = \lambda T^2(k) + (1 - \lambda)SPE(k), \text{with } 0 \leq l \leq 1 \quad (3)$$

An observation that associates with a change in the relationship between variables (i.e. outside the PCA model) will increase the *SPE*, whereas an observation inside the PCA model but distant from its center will manifest itself with an increase in the T^2. The parameter λ controls the contribution from extreme observations inside the PC subspace (e.g. uncommon operating conditions) over the contribution from observations outside the PC subspace (e.g. instrumentation drifts and erroneous measurements). The type of anomalies we are interested in are deviations from the principal modes of variation of the array of filters, this suggests a low value for λ. In the experiments λ was thus set to be equal to 0.1, which gives most of the weight to SPE while still retaining some information from T^2. In addition, a restrictive threshold J_t on the *J-statistic* has been set to identify only those observations that truly belong to the low-dimensional PCA model; J_t has been selected so that only 10% of the observations are retained (the 10-quantile). Such samples have been flagged as reliable and then used to learn the regression models.

Variable selection

Among all the candidate variables ($7 + 5 + 12$ for each filter), only a small but yet representative subset that truly contributes to a macroscopic characterization of the cell in terms of nitrogen removal has been identified. For the purpose, variable selection has been approached starting from the physical understanding of the nutrient removal process and has been further confirmed from the analysis of the correlation coefficients between the key process variables and the individual nitrate concentration in the filter. The choice of correlation as a measure of dependence is justified by the need of computationally light linear models to be implemented in the control system.

On the attempt of defining parsimonious models, with only a few representative secondary variables accessible to process experts, for each filter only six variables are selected as inputs for the sensors; namely, we selected influent and effluent nitrate-nitrogen (*I-NO*$_3$ and *E-NO*$_3$), influent wastewater and methanol flow rates (*Fi-QW*, the sum of *Fi-QW-1/2* and *Fi-QM*, the sum of *Fi-QM-1/2*), the head-loss (*Fi-HL*) in the filter and the effluent temperature (*E-T*).

Regression models

As already mentioned, simplicity is one of the main requirements to allow a direct implementation of the sensors in the plant's control system. Mainly for such a reason, classic linear models such as ordinary least squares regressions (OLSR) and partial least squares regression (PLSR) have been considered for the design of the soft-sensors. Due to the intrinsic nonlinearity of biological processes and for the sake of completeness, we also considered a nonlinear but yet computationally light regression method: local linear regression based on k-nearest neighbors (k-NN LLR).

OLSR is a classic regression method that learns a reconstruction of the functionality between the D-dimensional input measurements $\mathbf{x}(k)$ and the output measurements $y(k)$ by estimating the regression vector $\boldsymbol{\beta}$ $[\beta_1, \ldots, \beta_d, \ldots, \beta_D]'$ that parameterizes the linear model $y(k) = \boldsymbol{\beta}\mathbf{x}(k) + r(k)$, see Hastie *et al.* (2009) for reference. The estimation of $\boldsymbol{\beta}$ is based on a criterion that minimizes the residual sum of squares by globally fitting a D-dimensional hyper-plane over the available learning data. PLSR (Wold 1975) is also a global regression method that learns a linear model between the inputs and output but, instead of using the original variables, the fitting is iterative and based on their decomposition in latent variables. The decomposition in PLSR aims at maximizing simultaneously the variance in the inputs and the covariance between the inputs and the output. The PLSR model is also a parametric method with an additional (meta-) parameter, the number S of latent variables to be retained for prediction; in our experiments, S has been optimized using a standard resampling method for cross-validation, the Leave-One-Out, LOO (Hastie *et al.* 2009).

In their basic formulation, Local Linear Regression, LLR (Gupta *et al.* 2008) methods are nonlinear regressors based on the same principle of OLSR, the main difference being the approach used in the estimation of the regression coefficient. In LLR, instead of a global fitting of the least-squares hyper-plane, the fit is performed locally only in the neighborhood of the input measurement $\mathbf{x}(k)$ for which the output $y(k)$ is to be predicted. Several strategies for the definition of locality exist when LLR is used in soft-sensor development (Zhu *et al.* 2011); in k-nearest neighbor LLR (k-NN LLR) the neighborhood of the input observation $\mathbf{x}(k)$ is defined by K of its neighbors, according to a predefined metric; conventionally, the Euclidean distance is used. K is the (meta-) parameter of k-NN LLR models; usually, it is either fixed beforehand or cross-validated. In the experiments, K has been set to be equal to 10% of the total number of observations globally available for learning.

RESULTS

Table 2 presents the estimation results for the soft-sensors when predicting samples that correspond to normal operations. The results show that overall quality of the estimations obtained with the soft-sensors is always comparable with the resolution of the hardware sensors ($+/-0.1$ mg/L of NO_3-N) for Filters 2 to 9. As expected, the accuracy slightly degrades for Filters 1 and 10. Interestingly, it can be noticed that only minor improvements are obtained when linear OLSR and PLSR models are replaced by nonlinear k-NN LLR models. It can also be noticed how OLSR and PLSR always achieve identical results; this is easily explained by the fact that the number of latent variables cross-validated for PLSR is always equal to the number of original inputs used by OLSR. Therefore, in our experiments, PLS only performed a mere change of variables without losing any of the original information; the resulting PLSR regression models are thus identical to OLSR.

In order to illustrate the potentiality of the soft-sensors, we discuss three situations selected as representative of different conditions in the denitrifying post-filtration unit. In the first two cases, it is shown how the soft-sensors can be used to replace an out-of-order analyzer and how they can be used as a back-up system. The third case illustrates a situation where nonlinear models are beneficial.

Case 1: Temporary malfunction of the hardware sensor

The first example shows a temporary malfunction of the nitrate-nitrogen sensor in Filter 9 experienced during the autumn season in 2009, Figure 2. In particular, Figure 2(b) shows that the hardware sensor in Filter 9 is not returning any measurement and demonstrates how the missing measurements can be replaced by the soft-sensor estimates. The good quality of the estimates is confirmed by looking at the results obtained for Filter 3, in Figure 2(d), where the analyzer's resolution ($+-0.1$ mg/L) is reported as a grey

Table 2 | Estimation results in prediction reported as RMSE (mg/L of NO_3-N) for the ten filters

	Filter 1	Filter 2	Filter 3	Filter 4	Filter 5	Filter 6	Filter 7	Filter 8	Filter 9	Filter 10
OLSR	0.38	0.18	0.14	0.16	0.23	0.19	0.22	0.24	0.26	0.31
PLSR	0.38	0.18	0.14	0.16	0.23	0.19	0.22	0.24	0.26	0.31
k-NN LLR	0.46	0.20	0.15	0.16	0.23	0.18	0.23	0.24	0.26	0.26

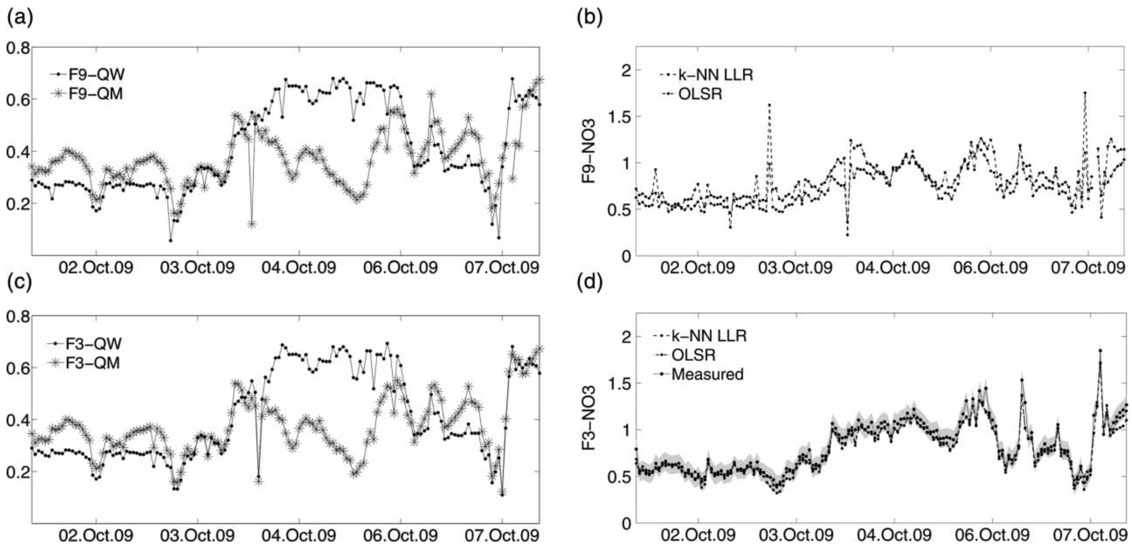

Figure 2 | Nitrate estimations when the filter sensor is properly working in Filter 3 (d) and absent in Filter 9 (b), under similar conditions in influent wastewater flow rates to Filter 3 (c) and Filter 9 (a).

band around the measurements. It is important to notice that Filters 3 and 9 are operated under very similar conditions, as exemplified in the diagrams by the influent wastewater and methanol flow rate, Figure 2(a) and (c).

Case 2: Erroneous measurement from the hardware sensor

The second case in Figure 3 refers to the situation in which erroneous measurements are returned by the hardware sensor in Filter 6. Also in this case, the soft-sensor is capable of recovering the nitrate concentration in the filter, Figure 3(a). This is confirmed by the comparison between the soft-sensor estimates and the measurements for a properly working analyzer; e.g. in Filter 3, Figure 3(b).

Case 3: Drastic weather

As a last example, a drastic weather event is considered, Figure 4. The snowmelt and rains in April 2010 caused an increase in the influent flow rate and associated decrease in the wastewater temperature, Figure 4(a), as well as variations in the NO_3-N concentration in the influent, Figure 4(c). The inherent nonlinear dependency of the process on temperature is underlined in the estimation results, reported for instance for Filters 4 and 5 in Figure 4(b) and (d). This variation is only partially recovered by the linear model, whereas *k*-NN LLR is able to identify it more accurately, thus confirming that, when the assumed linear dependence between the inputs and the output does not hold any more, sensible improvements can be achieved using nonlinear models.

DISCUSSION

The presented results show the possibility to implement the developed array of soft-sensors in the Viikinmäki plant and use it on-line for the estimation of the nitrate concentrations in the filters and, possibly, include the predictions in the methanol control loops. Before an actual implementation

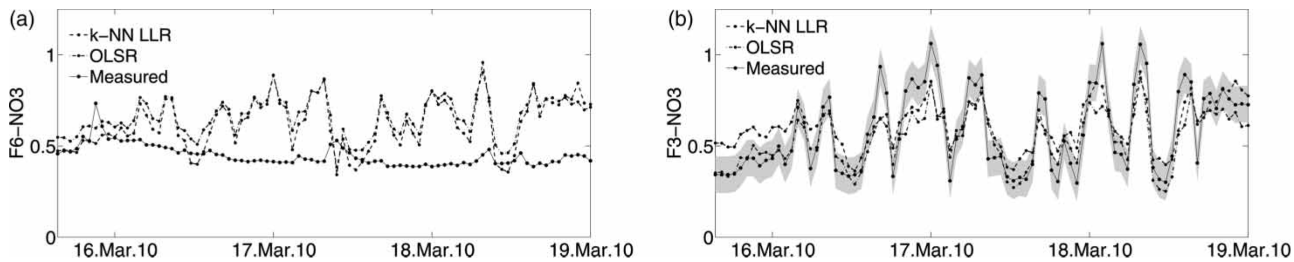

Figure 3 | Nitrate estimations when the filter sensor is not (a) and is properly (b) working.

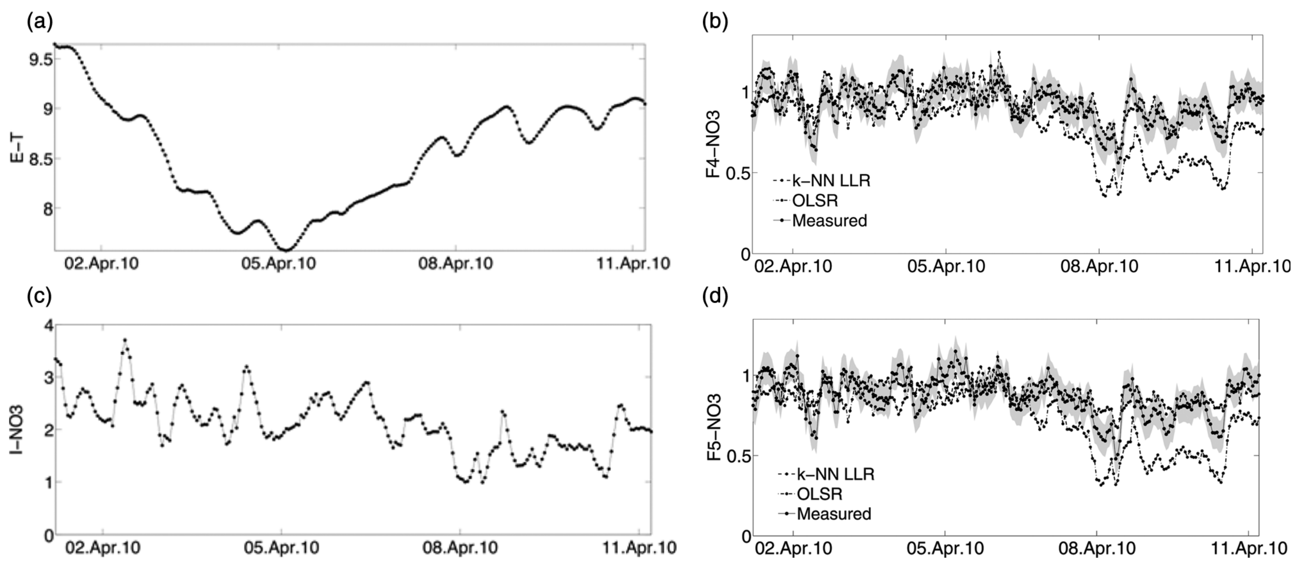

Figure 4 | Nitrate estimations in Filter 3 (b) and Filter 5 (d) during a snow-melting event that results in a temperature drop (a) and variations in the influent nitrate-nitrogen concentration (c).

in the plant's distributed control system, however, one important practical issue needs to be addressed: How to decide whether to use the software sensor or the hardware one? A consistent solution to this problem can be easily obtained by also including in the implementation the PCA models employed for the selection of reliable nitrate observations. The PCA models would, in fact, indicate in real-time the presence of anomalous nitrate measurements, and a threshold on the *J-statistics* could be used to trigger a switch between the analytical measurements and the soft-sensor estimates. Moreover, analogous PCA models could also be designed to detect anomalies among the input measurements and, when implemented in the plant's control system, thus also used to validate the quality of the estimates.

In addition, we would like to conclude the discussion on the development of the soft-sensors by pointing out that several improvements to our general methodology can be foreseen. Specifically, several alternatives for the design of regression models could be considered and also less pragmatic variable selection schemes could be used. In the linear framework, multivariable correlation coefficients could be used for selecting groups of input variables that best correlate with the output, according to a filtering approach. Or, alternatively, the variable selection step could be directly embedded in the learning of the regression model by using wrapper methods such as, the least absolute shrinkage and selection operator (LASSO) (Tibshirani 1996) and the least angle regression (LARS) (Efron *et al.* 2004). In a nonlinear framework, regression paradigms such as feedforward neural networks (Haykin 2009), support vector machine for regression (Cristianini & Shawe-Taylor 2000), Gaussian processes (Rasmussen & Williams 2006), coupled with filtering variable selection schemes based on mutual information (Rossi *et al.* 2006), could be employed. Such approaches are potentially valuable solutions that introduce only limited computational costs during the learning phase, but are easily processed once the models are already trained and implemented. A further improvement, valid for both linear and nonlinear regression methods, would consist of developing truly dynamic regression models that take into account also information on the temporal evolution of the process; this can be achieved, for example, by defining an input space that also includes several past measurements of the input variables and, possibly, also of the output variables. In general, the benefits of all the aforementioned sophistications are expected under special operating conditions such as strong deviations from linearity and for processes with dynamic characteristics much faster than the sampling rate. In our application, however, only minor improvements would be achievable, seeing that the accuracies obtained with the developed simple models are already comparable with the resolution of the analytical instrumentation. In that sense, simplicity is believed to be the right path to follow to advocate an ample diffusion of soft-sensing techniques in full-scale applications.

CONCLUSION

In this paper, results on the feasibility study for the development of an array of soft-sensors to estimate the nitrate-nitrogen concentration in the denitrifying post-filtration unit of the Viikinmäki WWTP are presented and discussed. From a practical point of view, the study has demonstrated the potential benefits to WWTP supervision and monitoring through the use of estimation devices. Such devices can be used as an inexpensive back-up system to conventional analytical equipment for both replacing out-of-order components and for validating existing field measurements. In the study, both linear and nonlinear estimation models have been considered. Based on our results, both typologies of models are capable of achieving accuracies that are overall comparable with the analytical instrumentation. As expected, nonlinear methods are capable of improving the estimation performances of linear methods, with such improvements mostly observable under ample variations in operating conditions but, achievable at the price of a slightly heavier computational cost.

ACKNOWLEDGEMENT

The authors acknowledge the Finnish '*Maa- ja vesiteknii-kan tuki ry*' for financial support.

REFERENCES

Aguado, D. & Rosen, C. 2008 Multivariate statistical monitoring of continuous wastewater treatment plants. *Engineering Applications of Artificial Intelligence* **21** (7), 1080–1091.

Avella, A. C., Görner, T., Yvon, J., Chappe, P., Guinot-Thomas, P. & de Donato, P. 2011 A combined approach for a better understanding of wastewater plants operation: statistical analysis of monitoring database and sludge physico-chemical characterization. *Water Research* **45** (3), 981–992.

Baggiani, F. & Marsili-Libelli, S. 2009 Real-time fault detection and isolation in biological wastewater treatment plants. *Water Science and Technology* **60** (11), 2849–2961.

Cristianini, N. & Shawe-Taylor, J. 2000 *Support Vector Machines*. Cambridge University Press, Cambridge.

Dochain, D. & Vanrolleghem, P. 2001 *Dynamical Modelling and Estimation in Wastewater Treatment Processes*. IWA Publishing, London, UK.

Efron, B., Hastie, T., Johnstone, I. & Tibshirani, R. 2004 Least angle regression. *Annals of Statistics* **32** (2), 407–499.

Gibert, K., Rodriguez-Silva, G. & Rodriguez-Roda, I. 2010 Knowledge discovery with clustering based on rules by states: a water treatment application. *Environmental Modelling and Software* **25** (6), 712–723.

Gupta, M. R., Garcia, E. K. & Chin, E. 2008 Adaptive local linear regression with application to printer color management. *IEEE Transactions on Image Processing* **17** (6), 936–945.

Hastie, T., Tibshirani, R. & Friedman, J. 2009 *Elements of Statistical Learning 2nd edition*. Springer, New York.

Haykin, S. 2009 *Neural Networks and Learning Machines*, 3rd edition. Prentice Hall, New York.

Jolliffe, I. T. 2002 *Principal Component Analysis*, 2nd edition. Springer, New York.

Kadlec, P., Gabrys, B. & Strandt, S. 2009 Data-driven soft sensors in the process industry. *Computers and Chemical Engineering* **33** (4), 795–814.

Kano, M. & Ogawa, M. 2010 The state of the art in chemical process control in Japan: good practice and questionnaire survey. *Journal of Process Control* **20**(ADCHEM 2009 Special number), 969–982.

Lee, M. W., Hong, S. H., Choi, H., Kim, J. H., Lee, D. S. & Park, J. M. 2008 Real-time remote monitoring of small-scaled biological wastewater treatment plants by multivariate statistical process control and network-based software sensors. *Process Biochemistry* **43** (10), 1107–1113.

Olsson, G., Nielsen, M. K., Yuan, Z., Lynggaard-Jensen, A. & Steyer, J. P. 2005 *Instrumentation, Control and Automation in Wastewater Treatment Systems*. Scientific and Technical Report No.15, IWA Publishing, London, UK.

Raich, A. & Cinar, A. 1996 Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. *AIChE Journal* **42** (4), 995–1009.

Rasmussen, C. E. & Williams, C. K. I. 2006 *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MS.

Rosen, C. 2001 A Chemometric Approach to Process Monitoring and Control – with Applications to Wastewater Treatment Operation. Ph.D. Thesis, Lund University of Technology. Lund, Sweden.

Rossi, F., Lendasse, A., François, D., Wertz, V. & Verleysen, M. 2006 Mutual information for the selection of relevant variables in spectrometric nonlinear modelling. *Chemometrics and Intelligent Laboratory Systems* **80** (2), 215–226.

Tibshirani, R. 1996 Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58** (1), 267–288.

Vanrolleghem, P. & Lee, D. 2003 On-line monitoring equipment for wastewater treatment processes: state of the art. *Water Science and Technology* **47** (2), 1–34.

Wold, H. 1975 Soft modeling by latent variables: The nonlinear iterative partial least squares (NIPALS) approach. In: *Perspective in Probability and Statistics, Papers in Honor of M. S. Barlett* (J. Gani, ed.). Academy Press, London.

Yoo, C., Villez, K., Hulle, W. V. & Vanrolleghem, P. 2008 Enhanced process monitoring for wastewater treatment systems. *Environmetrics* **19** (6), 602–617.

Zhu, Z., Corona, F., Lendasse, A., Baratti, R. & Romagnoli, J. A. 2011 Local linear regression for soft sensor design with application to an industrial deethanizer. In: *Proc. IFAC World Congress* International Federation of Automatic Control (IFAC), pp. 2839–2844.

First received 18 August 2011; accepted in revised form 12 December 2011