

SYSTEM AND METHOD FOR PREDICTING A PARAMETER OF WASTEWATER TREATMENT PROCESS

FIELD OF THE INVENTION

5

The present invention relates to a system and method for predicting a parameter of wastewater treatment process. In particular, the system and method is suited (but not limited to) the prediction of an effluent parameter in an Expanded Granular Sludge Bed (EGSB) process and will be described in such context.

10

DISCUSSION OF THE BACKGROUND ART

Throughout the specification, unless the context requires otherwise, the word “comprise” or variations such as “comprises” or “comprising”, will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

15

Furthermore, throughout the specification, unless the context requires otherwise, the word “include” or variations such as “includes” or “including”, will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

20

Anaerobic wastewater treatment processes are industrial processes that use microorganisms to breakdown contaminants in the wastewater under oxygen free environments. Typical variants of anaerobic wastewater processes used in the industry include anaerobic filters, down flow stationary fixed-film reactors, fluidized bed anaerobic reactors and Upflow Anaerobic Sludge Blanket reactors (UASB). The Expanded Granular Sludge Bed (EGSB) reactor is a specific application of the USAB with high wastewater treatment rate. In the EGSB process, wastewater passes through the reactor with high upward-flow velocity.

25

Due to complexity in the anaerobic wastewater treatment process which is highly sensitive to organic loading disturbance and dynamic changes in

30

concentration of substances in the influent wastewater, it is pertinent to ensure that the management of biomass generation is performed properly. Typically, biomass management is done reactively (i.e. after sudden changes in concentrations/disturbance) and this is undesirable because amongst other
5 problems, biomass requires extensive time for acclimatization to the influent wastewater. In cases where biomass inventory is low due to inability to react on time, process upsets and biomass loss may occur which result in significant downtime. Such downtime may further be compounded with associated financial costs due to inability to deliver, loss of time and resources. Therefore
10 there exists a pertinent need to move from a 'reactive' to a 'predictive' management of wastewater treatment process.

In view of the above, one way of predictive management is the use of mathematical models to predict the key composition qualities of treated wastewater. Such an approach reduces any hindrance to effective operation of
15 a wastewater treatment process and facilitates biomass inventory management. However, finding a suitable mathematical model to accurately predict the anaerobic wastewater treatment performance is a challenge due to its multivalent and non-linear nature, susceptibility to random noise and systemic error due to the use of measurement instruments, dynamic nature and time
20 delayed availability of process information. In addition, current mathematical modelling involving fundamental principles (for example mechanistic models) is inadequate or inefficient in modelling the dynamic complexities of anaerobic reactor systems. It is easily appreciated that as the complexity of a system increases, the computational time involved increases exponentially. This
25 makes mechanistic models unsuitable for 'near real-time' type prediction or where a reasonable turnabout time is expected.

Another approach is the use of soft sensors, and in particular statistical, data-driven soft sensors, which are built on recorded process historical data for the predicting effluent parameters such as Total Organic Carbon (TOC). While
30 such soft sensors have been adopted for some wastewater treatment

processes in the prediction and monitoring of effluent (treated water) parameters, there exists a need to further improve such soft sensors so as to reduce computational requirements and resources while simultaneously improving prediction accuracy.

- 5 There further appears no application of soft sensors to USAB or EGSB process for the prediction of effluent parameters such as TOC due to the complexity of these USAB or EGSB process and the extensive amount of parameters to be considered. In addition, there also exists a need to improve the accuracy and efficiency of such soft sensor prediction techniques by providing reliable
10 estimates for the aforementioned effluent parameters to evaluate the performance of EGSB reactor at a reasonable turnaround time (if not near real-time).

It is therefore an object of the invention to meet the above needs and alleviate the challenges at least in part.

15

SUMMARY OF THE INVENTION

- The invention involves the development of soft sensors to predict and monitor the operation of wastewater treatment operations, specifically in the area of Expanded Granular Sludge Bed (EGSB) reactors but not limited to such
20 reactors in the anaerobic treatment of industrial wastewater. In one embodiment, a soft sensor based on Partial Least Squares to predict TOC as a performance indicator to monitor EGSB reactor in a wastewater treatment plant (WWTP).

- In accordance with an aspect of the invention there is a method for predicting a
25 parameter of wastewater treatment process comprising the steps of:- (a) obtaining a dataset, the dataset comprising a plurality of process variables related to the parameter of wastewater treatment process to be predicted; (b) obtaining a predetermined number of measured values of the parameter to be predicted; (c) pre-processing the obtained dataset; the pre-processing step

includes aligning the process variables to account for any time-varying nature of the process variables and sorting the dataset into a set of inputs, the set of input comprising the dataset and the measured values of the parameter to be predicted; (d) obtaining a resultant output at each soft sensor; and (e) averaging
5 the resultant output across each of the plurality of soft sensor; the averaged resultant output being set as the final predicted parameter.

Preferably, the wastewater treatment process includes an expanded granular sludge bed (EGSB) reactor process and the parameter to be predicted is the effluent total organic carbon (TOC) of the EGSB reactor.

10 Preferably, the step of pre-processing includes applying linear interpolation or other interpolation methods such as polynomial and spline to the dataset where there are gaps detected in the dataset.

Preferably, the sorting of the dataset includes expressing the dataset in an auto-regression with exogenous inputs (ARX) structure.

15 Preferably, the measured values of the parameter to be predicted are passed through a noise reduction filter before step (d.) The noise reduction filter may preferably comprise a Savitzky–Golay filter and/or a Kalman filter.

Preferably, there comprise four soft sensors. In such instance preferably the output from each of the four soft sensors is further processed to account for
20 inaccuracies arising from dynamic nature of the wastewater treatment process before step (e).

Preferably, the soft sensors include artificial neural networks, support vector machines, Gaussian process regression.

In accordance with a second aspect of the invention there is a system for
25 predicting a parameter of wastewater treatment process comprising a primary set of measuring means arranged to obtain a dataset, the dataset comprising a plurality of process variables related to the parameter of wastewater treatment process to be predicted; a secondary set of measuring means arranged to

- obtain a predetermined number of measured values of the parameter to be predicted; a processor arranged to:- (i) receive the dataset and predetermined number of measured values from the first and second set of measuring means; the processor operable to pre-process the obtained dataset and sorting the dataset into a set of inputs; the set of input further comprising the dataset and the measured values of the parameter to be predicted; (ii) obtain a resultant output at each soft sensor; and (iii) average the resultant output across each of the plurality of soft sensor; the averaged resultant output being set as the final predicted parameter.
- 5
- 10 Preferably, the wastewater treatment process includes an expanded granular sludge bed (EGSB) reactor and the parameter to be predicted is an effluent parameter.
- Preferably, the effluent parameter includes but not limited to total organic carbon (TOC); chemical oxygen demand (COD); and biochemical oxygen demand (BOD) of the EGSB reactor.
- 15
- Preferably, the EGSB reactor comprises at least one equalization tank and at least one influent conditioning tank.
- Preferably, the primary set of measuring means comprises online and offline measurements.
- 20 Preferably, the secondary set of measuring means comprises offline measurements.
- Preferably, the process variables include level of wastewater in the at least one equalization tank; flow rates between equalization tank and influent conditioning tank, influent and effluent temperature; wastewater pH at equalization tank and conditioning tank.
- 25

BRIEF DETAILS OF THE DRAWINGS

In order that this invention may be more readily understood and put into practical effect, reference will now be made to the accompanying drawings, which illustrate preferred embodiments of the invention, and wherein:

- 5 Fig. 1 shows an expanded granular sludge bed (EGSB) reactor process;
- Fig. 2 is a flow chart of a method for predicting a parameter of wastewater treatment processes in accordance with an embodiment of the invention;
- Fig. 3 is a flow diagram of an algorithm adopted for developing a soft sensor for use in predicting a parameter of waste water treatment processes; and
- 10 Fig. 4 shows the result of prediction in accordance with the method for predicting a parameter of wastewater treatment processes.

Other arrangements of the invention are possible and, consequently, the accompanying drawings are not to be understood as superseding the generality of the description of the invention.

15

DESCRIPTION OF EMBODIMENTS OF THE INVENTION

In the description, it is to be appreciated that the term 'effluent water' refers to treated wastewater from the anaerobic/EGSB reactors. The treated wastewater may or may not be of sufficient quality to be discharged, i.e. the treated water
20 output from the anaerobic/EGSB reactor may need to undergo further treatment before it achieves sufficient quality.

In accordance with an embodiment of the invention there is a method 200 for predicting a parameter for a wastewater treatment process (WWTP). In particular, the method is suitably applied (but not limited) to an EGSB
25 plant/reactor and the parameter to be predicted is the total organic carbon (TOC) in the effluent of the EGSB reactor.

As an illustration and with reference to Fig. 1, a typical EGSB 10 system and process comprise an influent conditioning tank 20 and an EGSB reactor 30. In the EGSB reactor 30, granulated microorganisms are contacted against influent
30 wastewater from the influent conditioning tank 20 flowing at relatively high flow rates. Wastewater is continuously treated until the wastewater is adequately

treated and then they are discharged as treated water 40. The consortium of microorganisms in the EGSB process breaks down organic contaminants in the wastewater to form additional biomass (microorganisms), treated water 40 and energy in form of biogas 50, which includes methane and carbon dioxide (CO₂).

- 5 Various parameters (as illustrated in Table 1) of the EGSB reactor 30 may be used for the prediction of an effluent parameter.

The predicted TOC may be utilized for, inter alia, the monitoring of processing performance of the EGSB reactor.

With reference to Fig. 2, the method 200 comprises the following steps:

10

(a.) Obtaining a dataset (step 202) comprising a plurality of process variables related to the TOC in the effluent of the EGSB reactor. The plurality of process variables include, but are not limited to:- TOC in the influent; pH; flow rate; temperature etc. In an embodiment, twenty-two different process variables are selected. The twenty-two process variables are based on an EGSB reactor having one or more equalization tanks in addition to one or more influent conditioning tanks 20. The reactor further comprises a plurality of heat exchangers operable to cool down the reactors during operation.

15

The twenty two parameters are listed in Table 1 as follows:-

20

Table1. List of process variables

Tag No.	Process variables	Tag No.	Process variables
1	Equalization Tank level	12	pH of ww in EGSB Conditioning Tank B
2	Equalization Tank outlet temperature	13	Valve opening for flow control valve between Conditioning Tank A and EGSB A
3	Equalization Tank ww pH	14	Valve opening for flow control valve between Conditioning

			Tank B and EGSB B
4	Equalization Tank ww TOC	15	EGSB effluent temperature
5	Flowrate of cooling stream to feed ww heat exchangers	16	EGSB effluent pH
6	ww temperature at outlet of feed wastewater heat exchanger	17	Equalization Tank ww pH (lab measurement)
7	ww flowrate to EGSB conditioning tanks	18	Equalization Tank ww TOC (lab)
8	Control valve opening of ww flowrate to EGSB conditioning tanks	19	Equalization Tank ww Total Suspended Solid (lab)
9	Biogas pressure	20	Equalization Tank ww Volatile Fatty Acid (lab)
10	Feed flow valve opening to Conditioning Tank B	21	Equalization Tank ww alkalinity (lab)
11	pH of ww in EGSB Conditioning Tank A	22	Equalization Tank ww phenol (lab)

Note: wastewater is abbreviated as 'ww' in the above table.

The process variables may be obtained from historical samples or data in the operation of the EGSB plant.

- 5 (b.) In addition to obtaining a dataset of the process variables, offline (non real-time) measurements of effluent TOC corresponding to the twenty two process variables are also be obtained as starting TOC data (step 210).
- (c.) The obtained dataset is next pre-processed. The pre-processing comprises the following steps:-
 - 10 i. Alignment of the process variables (step 204). As the dataset of process variables may have different sampling frequency in seconds, minutes, hours, or

days, in order to obtain the time-varying characteristics, the time base reference of the process variables is predetermined at a value, e.g. 24 hour, and the hourly data of the twenty-two process variables of the preceding 24 hour data is used to match with the initial values of daily effluent TOC. As a result, the dimension of input variables is 528 (22 multiplied by 24).

- ii. Missing value handling step (step 206). It is to be appreciated that in the waste-water treatment process, the physical sensors (online for 24 hours 7 days) may fail to obtain all required data. In this regard, different online sensors may fail to collect the necessary data at different time periods. The missing value handling step 206 is added to handle such missing gaps in data collection and may handle the gaps via interpolation means. Interpolation means includes, for example, linear interpolation but may include other interpolation methods such as for example polynomial and spline, or other missing value handling approaches can also be implemented to handle missing values.
- iii. Sorting the input process variables and initial output process parameter TOC (step 208). The input process variables are further arranged into an autoregressive exogenous (ARX) structure to account for the large delay and large time-varying characteristics. This is to account for dynamic process changes, such as serious random interference, large delay, and large time-varying characteristics so as to improve the reliability of the predictive performance.

In order to account for the time-varying and time-delay characteristics, the input vector $\mathbf{x}$ for the ARX structure is mathematically expressed in equation (1) as follows.

$$\mathbf{x} = [y(t-1), \dots, y(t-L), \mathbf{u}(t)] \quad (1)$$

wherein observed past L outputs and past L control vectors. The number of orders in ARX model structure may be optimized using the Akaike information criterion (AIC) to achieve optimized modeling performance for dynamic modeling. According to the feature of process, $\mathbf{u}(t)$ is the arranged process variables with 528 dimensionality at time k, and $y(t-1), \dots, y(t-10)$ are the L= 10 TOC values before time t.

iv. Signal processing of the measured values of TOC for the removal of noise in data (step 212). The noisy offline measurements of TOC values arising from observed output data obtained from process and laboratory measurements may lead to difficulties in effectively capturing the trend of process dynamics.

- 5 Filtering noise in the measurements of TOC is beneficial for the improvement of predictive performance from the developed soft sensor. In particular, filtering or smoothing away noise would improve the accuracy of a dynamic model and allow the intrinsic trend of the process to be observed clearly.

Two filters are adopted to filter the noise, Savitzky-Golay ("Sav-Gol") filter and

- 10 Kalman filter (but not limited, other noise filter approaches, such as Fourier transform, wavelet, or Wiener filter, can also be implemented to filter noise).

The Sav-Gol filter is operable to smoothen the data and increase signal-to-noise ratio without greatly distorting the original data. In a sense, the Savitzky-Golay filter is based on local polynomial regression, which can be interpreted
15 mathematically that, for each point j with value y_j , a weighted sum of the neighboring values is calculated (a linear combination) and the number of neighbors and the degree of the polynomial control the strength of smoothing, as follows (step 214):

$$\hat{y}_j = \frac{1}{N} \sum_{h=-k}^k c_h y_{j+h} \quad (2)$$

- 20 where $\hat{y}_j$ is the smoothen or processed new 'de-noised' TOC value (of a smoothed curve or a derivative), N is a normalizing constant, k is the number of neighboring values at each side of j , hence the window size is $2k+1$, and c_h are the coefficients that depend on the used polynomial degree and on the goal (smoothing, first or second derivative). A window size of 3, polynomial order 2
25 and derivative 0 is adopted for the application.

The Kalman filter as an alternative is initially describing a recursive solution to the discrete-data linear filtering problem, given by linear stochastic difference equation with a measurement as follows (step 216):

$$x(k+1) = \mathbf{A}x(k) + \mathbf{B}u(k) + w(k) \quad (3)$$

$$y(k) = \mathbf{C}x(k) + v(k) \quad (4)$$

where w and v are uncorrelated, normally distributed white noise processes with covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively, and S is the covariance between $\mathbf{w}$ and $\mathbf{v}$. Kalman filter could be divided into two distinct phases: prediction and
5 update (filter) as follows:

$$\hat{x}(k+1|k) = \mathbf{A}x(k|k) + \mathbf{B}u(k) \quad (5)$$

$$\mathbf{P}(k+1|k) = \mathbf{A}\mathbf{P}(k|k)\mathbf{A}^T + \mathbf{Q}(k) \quad (6)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + \mathbf{K}g(k+1)(y(k) - \mathbf{C}\hat{x}(k+1|k)) \quad (7)$$

$$\mathbf{P}(k+1|k+1) = (\mathbf{I} - \mathbf{K}g(k+1)\mathbf{C})\mathbf{P}(k+1|k) \quad (8)$$

$$10 \quad \mathbf{K}g(k+1) = \mathbf{P}(k+1|k)\mathbf{C}^T(\mathbf{C}\mathbf{P}(k+1|k)\mathbf{C}^T + \mathbf{R}(k))^{-1} \quad (9)$$

where $\mathbf{K}g$ and $\mathbf{P}$ are Kalman gain and the covariance matrix of the state estimation error, respectively. The prediction equations 5 and 6 are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step. The filter equations 7 and 8
15 are responsible for the feedback—i.e. for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate. It is the one-step-ahead filter modified by the measurement. A simply linear Kalman filter may be adopted to filter the noisy process data, with $A = 1$, $B = 0$, $\mathbf{C} = \mathbf{I}$, $P(0|0) = 1$, and $x(0|0) = 0$.

20 (d.) The actual algorithm (i.e. soft sensors) for processing the input and measured TOC values is based on a partial least square (PLS) algorithm and a stacked PLS (SPLS) algorithm, the governing equations mathematically expressed in equations 10 to 18. The input is a vector comprising the process variables and measured TOC values; the output comprises the initial measured
25 effluent TOC values.

A general form of the PLS algorithm is described as follows:-

Suppose $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T \in \mathbf{R}^{n \times m}$ and $\mathbf{Y}(\mathbf{y}_1, \dots, \mathbf{y}_n)^T \in \mathbf{R}^{n \times p}$ are the input and output matrices, , they could be linearly connected as follows:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \quad (10)$$

where $\mathbf{B}$ is a matrix of regression coefficients, and the $\mathbf{E}$ is the residual matrix.

- 5 Rather than finding this relation directly, both $\mathbf{X}$ and $\mathbf{Y}$ are modeled by linear latent variables according to the PLS regression model, i.e.

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E}_X \quad (11)$$

and

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{E}_Y \quad (12)$$

- 10 with the error matrices $\mathbf{E}_X$ and $\mathbf{E}_Y$, where the matrices $\mathbf{T}$ and $\mathbf{U}$ (score matrices) as well as the matrices $\mathbf{P}$ and $\mathbf{Q}$ (loading matrices) have columns, with $a \leq \min(m, q, n)$ being the number of PLS components. Then the $\mathbf{x}$ - and $\mathbf{y}$ -scores are connected by the inner linear regression,

$$\mathbf{u}_1 = b_1 \mathbf{t}_1 + h_1 \quad (13)$$

- 15 with h_1 and b_1 being the residuals and regression vector, separately, and $\mathbf{t}_1$ and $\mathbf{u}_1$ being the first PLS score component, respectively. Subsequently, second PLS component could be generated based on decomposing of residual matrices of $\mathbf{E}_{X,1}$ and $\mathbf{E}_{Y,1}$ shown from following Equations (14) and (15), through the same procedure as that in generation of the first component.

$$20 \quad \mathbf{E}_{X,1} = \mathbf{X} - \mathbf{t}_1 \mathbf{p}_1^T \quad (14)$$

$$\mathbf{E}_{Y,1} = \mathbf{Y} - b_1 \mathbf{t}_1 \mathbf{q}_1^T \quad (15)$$

- After repeating “ a ” times, where “ a ” ranges from 1 to 25, (but not limited to this range. It is to be appreciated that the range used here for PLS model depends
25 on the number of variables (process variables) in the input data), which is

determined by cross-validation, 'a' PLS components would be generated, and in each repeat, score vectors $\mathbf{t}_j$ ($j = 1, \dots, a$) are normalized instead of $\mathbf{w}_j$ and $\mathbf{p}_j$, where $\mathbf{w}_j$ is the j th weight vector. Therefore, PLS algorithm could be presented as follows,

$$5 \quad \{\mathbf{X}, \mathbf{Y}\} \xrightarrow{\text{PLS}} \{\mathbf{T}, \mathbf{W}, \mathbf{P}, \mathbf{B}, \mathbf{Q}\} \quad (16)$$

where $\mathbf{T}$, $\mathbf{W}$, $\mathbf{P}$, $\mathbf{B}$, $\mathbf{Q}$, are scores, weights, loadings, regression coefficients, and loadings matrices. Once regression matrix is generated, new outputs could be predicted through Eq. (11) based on new inputs. The prediction of new samples can be obtained as

$$10 \quad \hat{\mathbf{Y}} = \mathbf{X}_{\text{new}} \mathbf{B} \quad (17)$$

PLS modeling. Because of ARX structure of inputs and the outputs filtered by two filters, for easy reference in the figures and the description, the PLS soft sensors may be referred to as "ARX-PLS-SG" and "ARX-PLS-K", respectively. When the new input is available, the predicted TOC can be generated from these two soft sensors.

The stacked PLS algorithm (SPLS) may be used as an alternative or in conjunction with PLS for modeling EGSB process. The SPLS algorithm was developed by Ni et al. in 2009 and was previously applied in a different field of spectral calibration and prediction.

20 In view of the fact that historical data used for building up ARX-PLS mode involves the $\mathbf{X}$ matrix including about 1000 samples and 538 variables (10 of them are measured effluent TOC values), not all local regions of variables contain the same information. Some are more informative, but others may be redundant data. However, the redundancy variables in the dataset may cripple or
25 lower the predictive performance of PLS model. So, in order to mitigate the effects, the calibration set of input $\mathbf{X}$ is split into n disjoint intervals $\mathbf{X}_k$ of equal width and n PLS models are developed between target property vector $\mathbf{y}$ and

each of the n intervals $\mathbf{X}_k$. Then n PLS interval models are weighted according to the model's predictive performance in cross-validation (seen in equation 18).

$$\hat{W} = ARGmin(y - \sum_{k=1}^n w_k \hat{y}_k) \quad (18)$$

Referring to Fig. 3, after the arrangement of ARX structure, the input has a dimensionality of 538 (including ten measured effluent TOC as starting outputs). SPLS is developed for effective extraction of useful information in the input by using weights (see equation 18). The 538 variables are split into n ($2 \leq n \leq 20$) consecutive regions. If $n = 4$, the 538 variables are split into 4 local regions. In each local one, the PLS model is built base on equations 10 to 17. The 4 local PLS models are fused by the weights in accordance with equation 18 where $n = 4$. The weights signify the importance of corresponding local model. Because n varies from 2 to 20, there are 19 SPLS models. The SPLS model with best performance is selected for prediction. Similar to PLS soft sensors, in this invention, SPLS soft sensors can be named as ARX-SPLS-SG and ARX-SPLS-K.

The pre-processed inputs are iterated in each of the labelled algorithm 'ARX-PLS-SG' (step 222); 'ARX-SPLS-SG' (step 224); 'ARX-PLS-K' (step 226); and 'ARX-SPLS-K' (step 228). Each of the four algorithms will generate the respective predicted TOC value.

To account for uncertainty of reliability of the wastewater treatment system and process, which is essentially a dynamic process which vary according to time, it is commonly possible that dynamic models, example the PLS with ARX, may model the processes with unacceptable accuracy. The dynamic models, therefore, need to be reliable and robustly adaptive. For this reason, the bias of dynamic models can be updated (S. Mu et al, 2006 and F. Ahmed et al, 2009) by incorporating a updating scheme to correct the predictive output of models through adding an offset smoother $bias(k)$ defined in Eq. (19) which weights and sums the current bias, $bias_0(k)$ defined in Eq. (20) and the entire offset at the latest point $bias(k-1)$, with weighting factor ω varying between 0 and 1, to

improve predictive accuracy of dynamic models. Note that the initial value of $bias(0)$ is 0.

$$bias(k) = \omega \times bias_0(k) + (1 - \omega) \times bias(k - 1) \quad (19)$$

$$bias_0(k) = y_{obs}(k - 1) - y_{mod}(k - 1) \quad (20)$$

- 5 The y_{obs} and y_{mod} are the observed target values (from offline laboratory) and the predicted values from dynamic models, i.e. PLS, respectively. The final output of dynamic models, $y_{cor}(k)$, then, can be corrected as follows:-

$$y_{cor}(k) = y_{mod}(k) + bias(k) \quad (21)$$

- 10 For the prediction from each soft sensor, the bias update is applied to get the updated predictive results through using equations from (19) to (21) with $\omega = 0.7$. Steps 232, 234, 236, 238 corresponds to the application of the bias update to the respective soft sensors 222, 224, 226 and 228.

- 15 Once the four soft sensor outputs have been obtained and have passed through the biasing update to account for dynamic changes, the results of the four predicted TOC output are 'fused' by averaging to generate the final predicted TOC value (step 240). The average of the four predicted TOC is set as the final predicted TOC value. It is to be appreciated that the average may be a simple average or a weighted
20 average giving higher weightage to certain soft sensors.

The embodiment will next be described in the context of its application to an EGSB plant to demonstrate the predictive performance.

- 25 In order to demonstrate the predictive performance of developed soft sensors in this invention, a historical data set included 1438 samples were extracted from an EGSB plant's record from January 2007 to December 2012 for development of the various soft sensors to teach or train the soft sensor model for use. For testing, 425 samples were then extracted from the plant's record from January 2013 to August 2014

(about one and half years) for the simulation of online prediction. It is to be appreciated that as the number of samples for teaching or training increases, the predictor output will improve.

- 5 In measuring the viability and feasibility of the method, the commonly used performance measures such as relative root mean squared error of prediction (RRMSEP), which compares the mean prediction of the model to the output of the process is utilized. The RRMSEP is expressed mathematically in equation 22 below.

$$10 \quad \text{RRMSEP} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2} \quad (22)$$

where y_i and $\hat{y}_i$ are the system's output and output of prediction from dynamic models in i^{th} step, respectively, and n is the number of samples in test data set for simulation of online prediction. The RRMSEP in equation (22) is used to evaluate the developed soft sensors in the
15 invention.

Table 2. The predictive performance from the developed soft sensors

Soft sensors	ARX-PLS-SG	ARX-SPLS-SG	ARX-PLS-K	ARX-SPLS-K	The fused
RRMSEP (%)	18.72	15.07	17.03	11.44	10.66

The first four methods in Table 2 were followed by implementation of
20 bias update to generate the predictive performance that shown in Table 2, which is corresponding to the flow chart in Fig. 2. Derived from Table 2, SPLS (Fig. 3) can more effectively extract useful information in variable-wise than PLS, thus lead to better performance. The final step of the invention (see Fig. 3) is the fusion of 4 developed soft sensors to
25 achieve better predictive performance. The prediction curve of this invention is depicted in Fig. 4.

Although the embodiments above are described in the context of TOC prediction in the EGSB reactor process, the method could be applied to other wastewater treatment processes such as activated sludge treatment, membrane filtration and reverse osmosis, for example.

In the case of activated sludge treatment, aerobic bacteria are utilized to breakdown contaminants present in sewage and water. Such aerobic bacteria rely on the use of dissolved oxygen in the wastewater to convert organic contaminants to additional biomass, carbon dioxide and water. Applying the method as described in the embodiments, the parameter to be predicted may be the contaminant removal efficiency based on input process parameters such as influent waste flow rate, concentration, dissolved oxygen level and pH.

In the case of membrane filtration and reverse osmosis, which are physical treatment methods for water that utilize membrane filters to remove contaminants from water; the method described in the embodiment could be utilized to improve operations of the membrane process by providing prediction of membrane performance based on related process variables such as influent quality and transmembrane pressure.

The described embodiments are advantageous in at least the following ways:-

- The developed soft sensor in this invention has the capability to handle missing values, handling large number of process parameters, accounting for the large delay and large time-varying characteristics, compensating the uncertainty of the system by bias update, filtering the noise in the measurements, more effective information extraction by SPLS, and better predictive performance by fusing of 4 soft sensors. Because the PLS was employed in this invention to model the wastewater treatment processes, the prediction of effluent qualities and the fault detection and isolation may be achieved simultaneously.
- This invention provides two levels of model fusion to achieve better prediction of TOC. The first level of model fusion is fusing a few PLS models into a Stacked PLS soft sensor to achieve better prediction than

original PLS model. The second level is at a higher level of model fusion, which can fuse either the models from the same approach (several PLS models) or the models from different approaches (fuse PLS and SPLS or any other regression methods). For model fusion, fusing more models achieves better prediction in general.

It is to be understood that the above embodiments have been provided only by way of exemplification of this invention, and that further modifications and improvements thereto, as would be apparent to persons skilled in the relevant art, are deemed to fall within the broad scope and ambit of the present invention described herein. In particular:-

- More or less process variables may be used as inputs to the soft sensors. For example, more than twenty-two process variables may be used to achieve a better prediction taking into account more process variables where computational processing time is not as critical.
- In these described embodiments, there comprises four soft sensors and the fusion of the described four soft sensors is better than fusing only two SPLS soft sensors, for example. The number of soft sensors for fusion is not limited and it depends on available soft sensors but in general, fusion of more soft sensors.
- It is to be appreciated that a variety of soft sensors may be used for fusion. Other types of soft sensors include artificial neural network, support vector machine, Gaussian process regression, etc.
- Although the aforementioned embodiment(s) has described the parameter to be predicted in the context of the effluent TOC parameter, it is to be appreciated that other effluent parameters including but not limited to chemical oxygen demand (COD), biochemical oxygen demand (BOD).
- Although examples of the wastewater treatment process described in the embodiment(s) relate to an EGSB type wastewater treatment process, it

is to be appreciated by a skilled person that the method and system may be applied for the prediction of parameter(s) for other anaerobic wastewater treatment processes to improve and enhance accuracy of prediction. In particular, the use of a fusion of soft sensors is likely to
5 enhance and improve prediction performance for a variety of applications.

It would be further appreciated that although the invention covers individual embodiments, it also includes combinations of the embodiments discussed. Therefore, features described in one embodiment not being mutually exclusive
10 to a feature described in another embodiment may be combined to form yet further embodiments of the invention.

CLAIMS

1. A method for predicting a parameter of wastewater treatment process comprising the steps of:-

5 (a) obtaining from a first set of measuring means a dataset, the dataset comprising a plurality of process variables related to the parameter of wastewater treatment process to be predicted;

(b) obtaining from a second set of measuring means a predetermined number of measured values of the parameter to be predicted;

10 (c) pre-processing the obtained dataset using a processor; the pre-processing step includes aligning the process variables to account for any time-varying nature of the process variables and sorting the dataset into a set of inputs, the set of input comprising the dataset and the measured values of the parameter to be predicted;

(d) obtaining resultant outputs from a plurality of soft sensors; and

15 (e) averaging the resultant output across each of the plurality of soft sensor; the averaged resultant output being set as the final predicted parameter.

2. A method according to claim 1, wherein the wastewater treatment process includes an expanded granular sludge bed (EGSB) reactor and the
20 parameter to be predicted is the effluent total organic carbon (TOC) of the EGSB reactor.

3. A method according to claims 1 or 2, wherein the step of pre-processing includes applying linear interpolation or other interpolation methods such as polynomial and spline to the dataset where there are gaps detected in the
25 dataset.

4. A method according to claim 3, wherein the sorting of the dataset includes expressing the dataset in an auto-regression with exogenous inputs (ARX) structure.
5. A method according to any one of the preceding claims, wherein the measured values of the parameter to be predicted are passed through a noise reduction filter before step (d.).
6. A method according to claim 5, wherein the noise reduction filter comprises a Savitzky–Golay filter and/or a Kalman filter.
7. A method according to any one of the preceding claims, wherein there comprise four soft sensors.
8. A method according to claim 7, wherein the output from each of the four soft sensors is further processed to account for inaccuracies arising from dynamic nature of the wastewater treatment process before step (e).
9. A method according to any one of the preceding claims, wherein the soft sensors include artificial neural networks, support vector machines, Gaussian process regression.
10. A method according to any one of the preceding claims, wherein the plurality of soft sensors include a partial least square (PLS) based soft sensor and a stacked PLS (SPLS) soft sensor.
11. A method according to claim 10, wherein there are four soft sensors.
12. A method according to claim 11, wherein there are two PLS based soft sensors and two SPLS based soft sensors.
13. A system for predicting a parameter of wastewater treatment process comprising:-
first set of measuring means arranged to obtain a dataset, the dataset comprising a plurality of process variables related to the parameter of wastewater treatment process to be predicted;

second set of measuring means arranged to obtain a predetermined number of measured values of the parameter to be predicted;

a processor arranged to:-

- 5 (i) receive the dataset and predetermined number of measured values from the first and second set of measuring means; the processor

operable to pre-process the obtained dataset and sorting the dataset into a set of inputs; the set of input further comprising the dataset and the measured values of the parameter to be predicted;

(ii) obtain resultant outputs at a plurality of soft sensors; and

5 (iii) average the resultant output across each of the plurality of soft sensor; the averaged resultant output being set as the final predicted parameter.

14. A system according to claim 13, wherein the wastewater treatment process includes an expanded granular sludge bed (EGSB) reactor and the
10 parameter to be predicted is an effluent parameter

15. A system according to claim 14, wherein the effluent parameter includes but not limited to total organic carbon (TOC); chemical oxygen demand (COD); and biochemical oxygen demand (BOD) of the EGSB reactor.

16. A system according to claim 14 or 15, wherein the EGSB reactor
15 comprises at least one equalization tank and at least one influent conditioning tank.

17. A system according to any one of claims 14 to 16, wherein the primary set of measuring means comprises online and offline measurements.

18. A system according to any one of claims 14 to 16, wherein the secondary
20 set of measuring means comprises offline measurements.

19. A system according to claim 16, wherein the process variables include level of wastewater in the at least one equalization tank; flow rates between equalization tank and influent conditioning tank, influent and effluent temperature; wastewater pH at equalization tank and conditioning tank.

20. A system according to any one of claims 13 to 19, wherein the plurality of soft sensors include a partial least square (PLS) based soft sensor and a stacked PLS (SPLS) soft sensor.
21. A system according to claim 20, wherein there are four soft sensors.
- 5 22. A system according to claim 21, wherein there are two PLS based soft sensors and two SPLS based soft sensors.