Estimation of COD removal via soft sensor

S. Marsili-Libelli*

*Department of Information Engineering, University of Florence, Florence, Italy
(stefano.marsililibelli@unifi.it)

Abstract: In conventional wastewater treatment one of the main problems confronting the plant manager is the estimation of the organic carbon removal efficiency, from which a mass balance can be computed. This paper presents a simple, albeit approximate, solution to this problem via a soft sensor approach. The basic idea is to apply a set-point control to the dissolved oxygen in the aerated compartment, so that the air supply is directly proportional to the removed rapidly degradable COD and to use the air supply information as an indirect estimation of the metabolized fraction of the biodegradable carbonaceous load. The paper analyses the performance of the soft sensor in a variety of operational situations.

Keywords: Wastewater Treatment Operation; Soft Sensors; PID control; Benchmark simulation models.

ENGINEERING PRINCIPLES

Respirometric techniques have been used for decades to infer the degree of microbial activity in a biological wastewater treatment plant (Spanjers and Vanrolleghem, 1995; Marsili-Libelli and Vaggi, 1997; Brouwer et al., 1998; Spérandio and Etienne, 2000). The relationship between the air flow rate and the biodegradable BOD abatement has been can be influenced by the aerator characteristics and the state of the biomass, but in any case it can be expressed by a linear relationship that our soft-sensor is meant to approximate. Therefore, the implementation of the soft sensor reduces to the design of a linear observer relating the input-output COD difference (ΔCOD) to the air flow rate (U_a)

$$\Delta COD = COD_{in}(t) - COD_{out}(t - \vartheta_H) = p_1 \times U_a + p_2 \tag{1}$$

where the output COD is delayed by the hydraulic retention time ϑ_H. So the input/output data to feed into the estimator are

$$\begin{aligned} &Input: && U_a \\ &Output: && \Delta COD = COD_{in}(t) - COD_{out}(t - \vartheta) \end{aligned} \tag{2}$$

The practical scheme of the plant equipped with the COD soft sensor is shown in **Errore. L'origine riferimento non è stata trovata.**.

Of the several options available, the most promising proved to be the fuzzy approach, in which the relationship (1) was replaced by a fuzzy inference system.

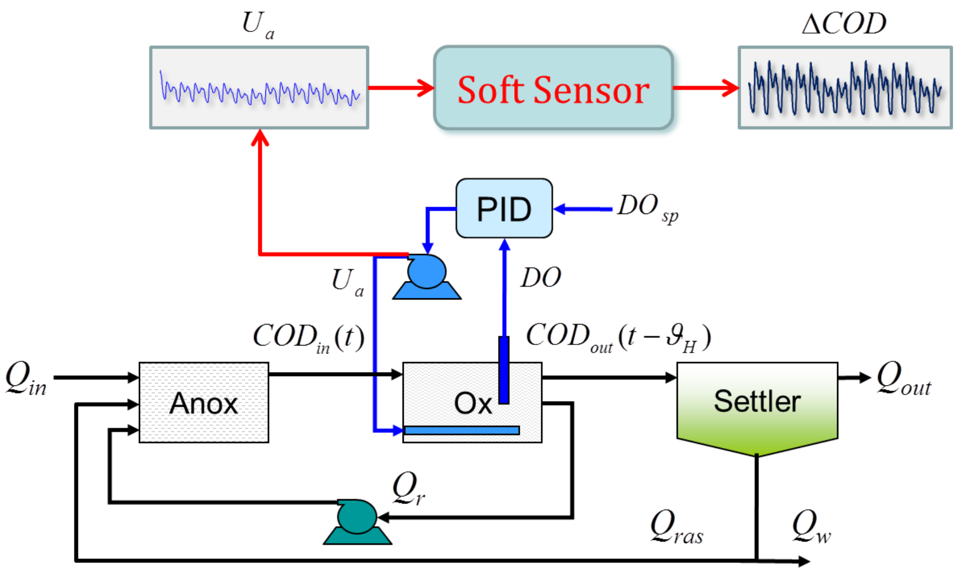

Figure 1. Overall scheme of the CAS process with a PID controller to keep the DO level constant. The soft sensor is shown in the upper part.

A FUZZY VERSION OF THE SOFT SENSOR

The previous considerations demonstrate that the linear approximation, though theoretically attractive, is not enough accurate when tested with a benchmark-level model, let alone field data. This may be due to the many components forming the COD and the numerous kinetics governing its transformation. For this reason a fuzzy extension of eq. (1) is introduced in the form of a series of local linear approximations, each approximating the $\Delta COD/U_a$ pseudo-linear relationship in a specific portion of the operating range. The fuzzy CSS estimator can then be formalised by the following set of rules

$$R_i : \; if \; U_a \subset C_i \; then \; \Delta COD_i = a_i U_a + b_i \quad (i = 1,\ldots,c)$$

$$\Delta C\hat{O}D = \frac{\sum_{i=1}^{c} \mu_i \times \Delta COD_i}{\sum_{i=1}^{c} \mu_i} \quad , \qquad (3)$$

where c is the number of clusters and the antecedent expression $U_a \subset C_i$ should be interpreted as “the degree of membership (μ_i) of U_a to cluster C_i”. This will weights the consequent linear approximation in eq. (3) $\Delta COD_i = a_i U_a + b_i$. The COD estimate is then given by the weighted sum of all the local approximations. The antecedent clusters (C_i) are obtained by clustering the U_a data by Fuzzy C-Means (FCM) clustering (Bezdek, 1981) and have the shapes shown in Figure 2, while the consequents are the linear segments shown in Figure 3 over the background of the training data.

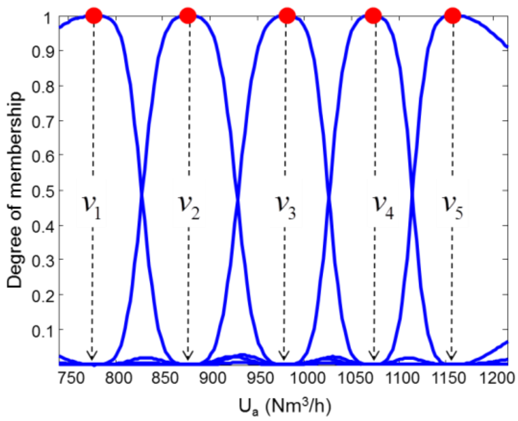

Figure 2. Membership functions generated by eq. (1.10).

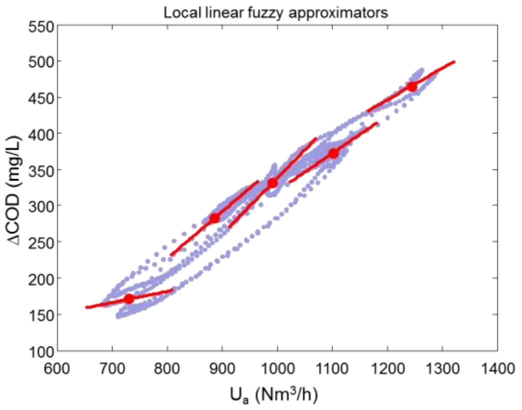

Figure 3. The consequents of the eq. (3) provide the local linear approximations to the observed behaviour.

The model identification was performed in two steps: first the antecedent membership functions are numerically identified via clustering, producing the partitioning of Figure 2, then the coefficients of the consequent linear approximations are identified in order to minimize the following error functional

$$SSE = \frac{1}{N}\sum_{k=1}^{N}\left(\Delta COD_k^{exp} - \Delta C\hat{O}D_k\right)^2, \tag{4}$$

where $\left\{\Delta COD_k^{exp}\ k=1,\ldots,N\right\}$ is the set of experimental COD measurements produced by the benchmark simulation. Of each data subsets, the first was used to calibrate the SS and the second for validation, after splitting the data between weekdays and weekends, as shown in Figure 4, given the widely differing characteristics of these two periods of time.

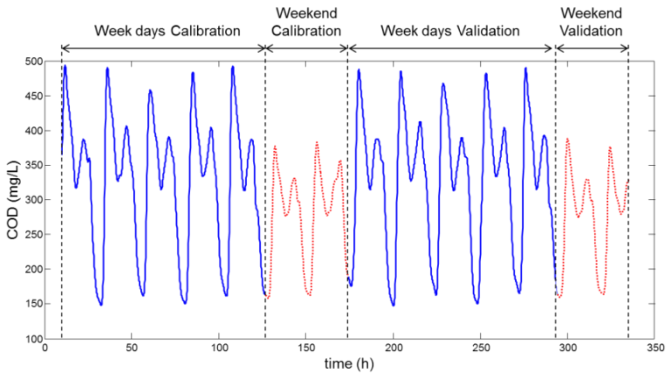

Figure 4. Benchmark data splitting between calibration and validation subsets.

The overall CSS scheme is summarized in Figure 5.

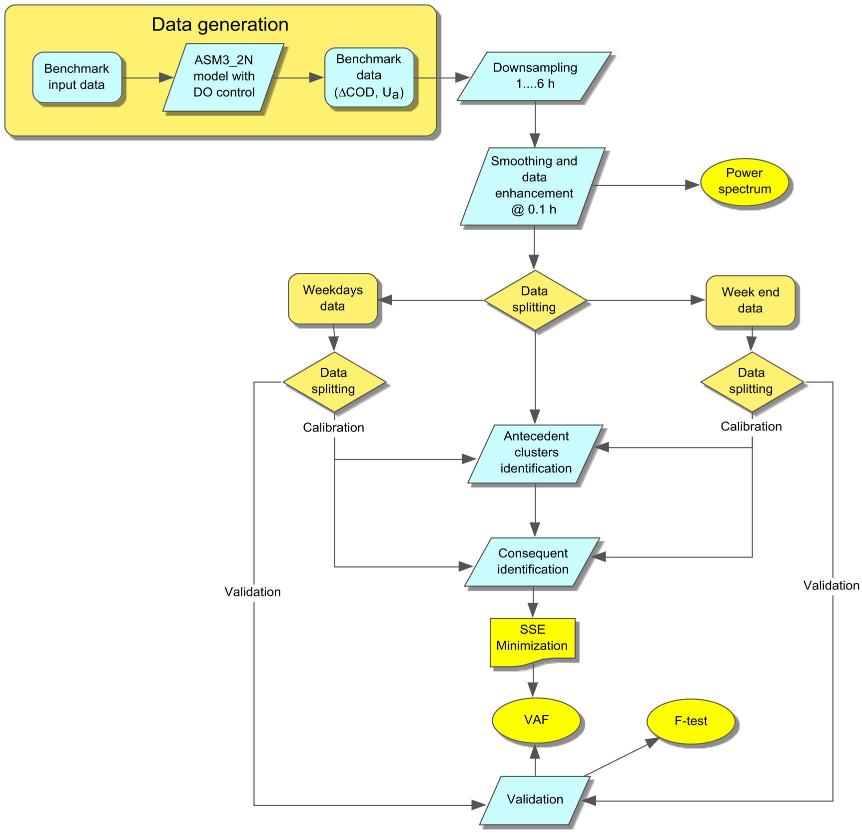

Figure 5. Overall scheme of soft sensor identification, including data preparation and.

As an example of the Soft Sensor performance, Figure 6 shows the calibration and validation data for synthetic data sampled at 1 h intervals.

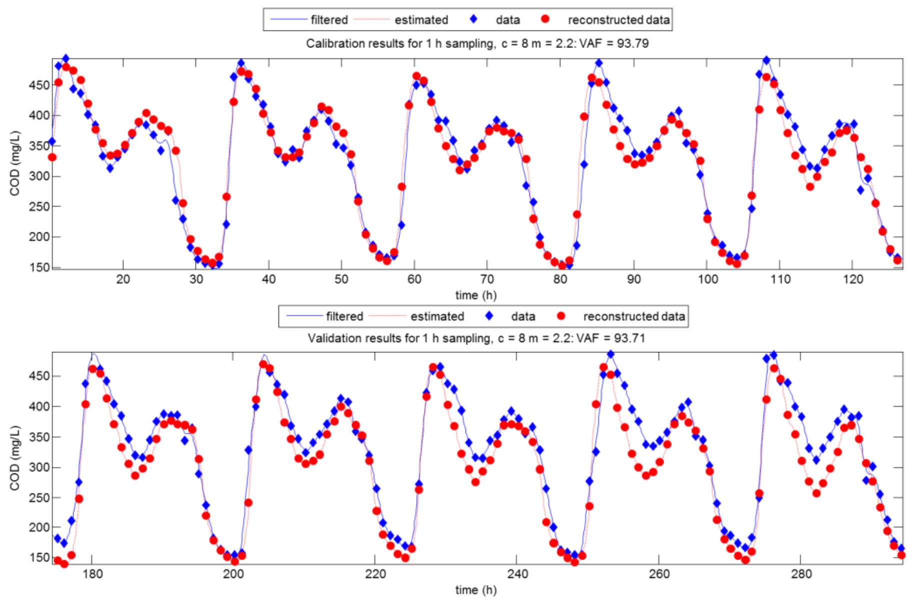

Figure 6. Calibration (above) and validation (below) of the weekdays portion of the synthetic data sampled at 1 h interval and enhanced at 0.1 h.

Validation of the soft sensor performance

In addition to recording the SSE and VAF for both weekdays and weekends data the soft sensor performance was validated by analysing the regression line fitted to the scatter plot comparing the model predictions to the calibration and validation by applying the F statistics to the regression line between model and data (Haefner, 2005) Figure 7 shows the regression lines for the calibration and validation data in the case of 1 h sampled data. In both cases the F values indicate that the null hypothesis (valid model) cannot be rejected at the 95% confidence level. The final CSS performance is summarized in Table 1.

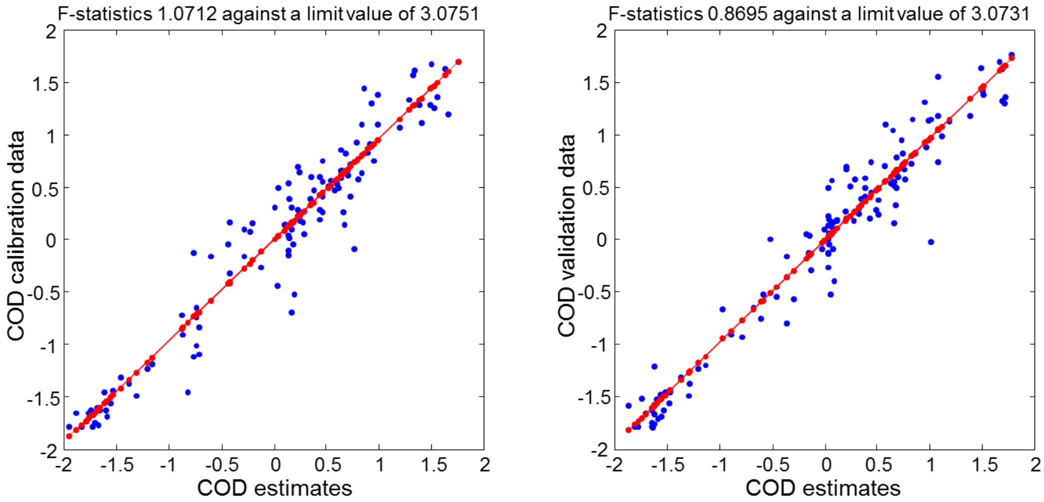

Figure 7. Regression lines and F statistics for the calibration and validation data of the soft sensor trained with the 1 h sampled data.

Table 1. Performance metrics of the optimally identified soft sensors for the weekdays and weekend data subset for a varying sampling interval T_s.

T_s		1		2		3		4		5		6	
Metrics		F	VAF	F	VAF	F	VAF	F	VAF	F	VAF	F	VAF
Weekdays	cal	1.0712	93.78	0.3910	94.64	0.3427	92.76	0.2204	93.81	0.1805	96.10	0.1954	91.56
	val	0.8695	93.71	0.4406	93.16	0.3763	91.98	0.1805	94.92	0.1554	94.35	0.3200	76.21
Weekend	cal	0.3927	93.22	0.2110	92.61	0.2284	87.37	0.2021	85.05	0.1575	82.94	0.1913	76.53
	val	0.2074	95.72	0.1019	95.51	0.0999	92.72	0.0520	94.32	0.0209	97.18	0.2230	67.52

REFERENCES

Bezdek, J. C., 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, pp. 256.

Brouwer, H., Klapwijk, A., and Keesman, K. J., 1998. Identification of activated sludge characteristics using respirometric batch-experiments. *Water research*, **32**, 1240–1254.

Copp, J. B., 2002. *The COST Simulation Benchmark: Description and Simulator Manual*, Office for Official Publications of the European Community, Luxembourg, pp. 154.

Dainotto, A., Barni, G., Giaccherini, F., Magnolfi, V., Marsili-Libelli, S., and Simonetti, I., 2012. A multilevel coordinated control strategy for energy conservation in wastewater treatment plants. In: *PID '12, IFAC Conference on Advances in PID Control*. Brescia.

Haefner, J. W., 2005. *Modeling Biological Systems, principles and applications*, Springer Science & Business Media, New York, pp. 475.

Iacopozzi, I., Innocenti, V., Marsili-Libelli, S., and Giusti, E., 2007. A modified Activated Sludge Model No. 3 (ASM3) with two-step nitrification-denitrification. *Environmental Modelling and Software*, **22**, 847–861.

Marsili-Libelli, S. and Vaggi, A., 1997. Estimation of respirometric activities in bioprocesses. *Journal of Biotechnology*, **52**, 181–192.

Spanjers, H. and Vanrolleghem, P. A., 1995. Respirometry as a tool for rapid characterization of wastewater and activated sludge. *Water Science & Technology*, **31**, 105–114.

Spérandio, M. and Etienne, P., 2000. Estimation of wastewater biodegradable COD fractions by combining respirometric experiments in various So/Xo ratios. *Water research*, **34**, 1233–1246.

Takács, I., Patry, G. G., and Nolasco, D., 1991. A dynamic model of the clarification thickening process. *Water Research*, **25**, 1263–1271.