

A novelty detection approach for detecting faulty batches in a photo-Fenton process

I. Monroy^a, E. Yamal^a, G. Escudero^b, M. Pérez-Moya^a and M. Graells^a.

^aChemical Engineering Department, ^bSoftware Department. Universitat Politècnica de Catalunya, EUETIB. Comte d'Urgell 187, Barcelona 08036, Spain.

Abstract

A novelty detection approach for detecting novel faults has been developed and applied to a photo-Fenton process in a fully monitored pilot plant. The proposed approach consists of two stages: Fault Detection or binary classification stage, and Fault Diagnosis (FD) or multi-class classification stage. Batches under nominal operating conditions, where coffee samples are degraded, are defined according to an experimental design and three possible faults are found in the process. Experimental batches under such normal and abnormal conditions are used to construct the classification models and therefore, the novelty detection approach. Two faulty batches, not learned by the models, were used to validate the approach. The successful results obtained encourage performing potential models for other pollutants that allow further detecting novel contaminants in wastewater treatments.

Keywords: Novelty Detection, Fault Diagnosis, Advanced Oxidation Process, photo-Fenton process.

1. Introduction

Fault Diagnosis (FD) in chemical processes based on historic-data models has received increasing attention in the last decade. Many classification algorithms from the Artificial Intelligence (AI) area have been successfully used as diagnosis methods. However, the diagnosis of “novel” faults is a difficulty to overcome and has been barely addressed in this area despite its importance (Burbeck and Nadjm-Tehrani, 2007; Ha An et al., 2011). The detection of faults not learned with the diagnosis methods is called Novelty Detection (ND) or Anomaly Detection (AD) (Patcha and Park, 2007).

Photo-Fenton process is an Advanced Oxidation Process (AOP) used for the remediation of wastewaters containing recalcitrant contaminants. It consists of the in-situ generation of hydroxyl radicals ($\cdot\text{OH}$) from the reaction between H_2O_2 and an iron salt in presence of UV radiation. The final remediation achieved depends extremely on the process variables, especially on those concerning to dosing reagents and to light intensity and wavelength (Pignatello, 2006).

In photo-Fenton and photolysis reactions, Fe^{3+} is excited when it is irradiated between 300 and 450 nm wavelengths, which increases the reaction velocity of Fenton process and produces more hydroxyl radicals. The reaction between Fe^{2+} and Fe^{3+} with H_2O_2 allows hydroxyl radical generation, which mineralizes organic matter into CO_2 , H_2O and inorganic acids (Tokumura et al., 2008). Moreover, it has been reported that the ratio between $\text{Fe}^{2+}/\text{H}_2\text{O}_2$ and pH affect reaction rate (Gulkaya et al., 2006).

This work presents an ND approach using Support Vector Machines (SVM) as classification algorithm and it is applied to a fully-monitored photo-Fenton pilot plant, operating batchwise. This approach combines simultaneously a detection or binary classification stage with a diagnosis or multi-class classification stage. In both stages, it

is required previously data representation and improvement steps before applying properly the classifier algorithm.

2. Materials and Methods

2.1. Case study

The pilot plant, shown in Figure 1, is composed of two reactors and a pump system of 0.5 HP that allows flow recirculation and mixture homogenization. Both reactors can work in series or independently. The main reactor is a 12L jacketed glass reactor equipped with a medium pressure Hg lamp of 700W in a glass lampholder, capable of irradiating between 390 and 410nm (photo-Fenton process). In contrast, the second reactor is a 2L tubular reactor with a 55W low pressure Hg lamp irradiating at 254nm in a quartz lampholder (photolysis process). After the tubular reactor, there is a motorized valve that deflects a small part of the flow to the measurement line.

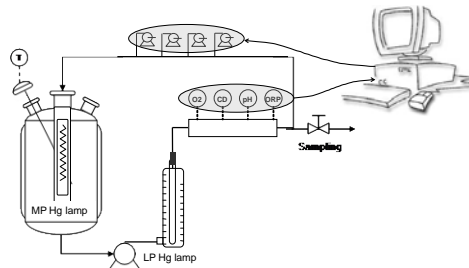


Figure 1: Pilot plant schematic representation

Thus, the plant is monitored with an automatic control and a data acquisition system in which eight variables are on-line registered: dissolved oxygen (DO), redox potential (ORP), pH, conductivity, main reactor temperature (T1), on-line temperature (T2), recirculation flow (Q1) and measurement line flow (Q2). Measurements are sampled each second. Recirculation flow is controlled by the pump and the measurement line flow and pH are controlled with PID controllers. The flexible pilot plant could then be configured in several ways affecting volume and irradiation source as indicated in Table 1. Moreover, different contaminants are also possible to treat with photo-Fenton.

Table 1. Plant configurations considered in the pilot plant

Class of irradiation	Configuration medium/low P Hg lamp
A	OFF/OFF
B	ON/OFF
C	OFF/ON
D	ON/ON

Furthermore, total organic carbon (TOC) and H_2O_2 concentrations are off-line measured at different time scales (each 10 or 15 minutes) in photo-Fenton reactions.

2.2. Novelty detection approach

The ND approach consists of two stages: binary and multi-class classification stages. First, both classifiers are set up by constructing and validating the classification models for each registered fault. The detection stage joins the faulty observations in a positive class and nominal data in a negative class. On the contrary, in the diagnosis stage, same training data is just classified according to the corresponding fault, thus leading to as

many diagnosis models as faults. Therefore, a novel fault is detected when it is detected (stage 1) and there is a lack of positive diagnosis of all the known faults (stage 2). In order to apply the ND approach to the photo-Fenton plant, the nominal operating conditions (NOC) are fixed according to Table 2, as well as the experimental design involving the two factors governing the H₂O₂ dosage: t_{ini} and y_0 . Equivalent initial H₂O₂ concentration ($C_{eq,\infty}^{H_2O_2}$) can be determined according to the next equation:

$$C_{eq,\infty}^{H_2O_2} = \frac{Q^{H_2O_2}}{V_R}$$

where $Q^{H_2O_2}$ is the total quantity of hydrogen peroxide, used for each batch and V_R is the reactor operation volume. In addition, dosage protocol is defined according to the next equations for the dosage variable $y(t)$:

$$y(t) = \begin{cases} 0 & \text{if } t < 0 \\ y_0 & \text{if } 0 \leq t < t_{ini} \\ y_0 + \left(\frac{1-y_0}{\Delta t_{add}} \right) (t - t_{ini}) & \text{if } t_{ini} \leq t < t_{ini} + \Delta t_{add} \\ y_0 & \text{if } t_{ini} + \Delta t_{add} \leq t < TS \end{cases}$$

Table 2. Variables and parameters values in the photo-Fenton process at NOC

Variable/Parameter	Value / Arrangement	Units
Contaminant (CTI)	Coffee	-
Irradiation	C	-
Contaminant initial concentration (CTI_0)	300.0	mgL ⁻¹
Initial Fe ²⁺ concentration ($C_0^{Fe(II)}$)	10.0	mgL ⁻¹
Equivalent initial H ₂ O ₂ concentration ($C_{eq,\infty}^{H_2O_2}$)	500.0	mgL ⁻¹
H ₂ O ₂ initial dosage time (t_{ini})	0 – 36.2	min
H ₂ O ₂ initial percentage of the total dosage (y_0)	5.9– 34.1	%
Reactor operation volume (V_R)	8.0	L
pH	3.0	-
Dosage time (Δt_{add})	60.0	min
Recirculation flow (Q_I)	11.3	Lmin ⁻¹
Pump work (PW)	75.0	%
Batch duration (TS)	120	min

Records from three different situations out of the experimental design and thus of the NOC were available and considered as faults. These three faults are:

- $C_{eq,\infty}^{H_2O_2}$ higher than 500 mgL⁻¹, up to 1500 mgL⁻¹ (fault 1)
- Change of configuration of the irradiation source from C to B (fault 2)
- Different H₂O₂ dosage protocol out of the experimental design (fault 3). Specifically, the y_0 range varies from 70 to 100% and in 0%. For this last case, Δt_{add} is 120 min.

These three faults were used for constructing the detection and diagnosis models required in both stages of the ND approach. Moreover, two different scenarios were considered as novel faults so that the approach could be validated. These novel faults are $C_0^{Fe(II)}$ higher than 10ppm and configuration A in the irradiation source (lack of

irradiation). The performance in both stages of the methodology is measured in terms of the *F1* score (Van Rijsbergen, 1979).

3. Results and Discussions

A set of twenty experimental batches under NOC and 10 batches per fault were used for constructing the models due to the lack of more available batches and the time consumed for their generation. The final sampling time considered for the models construction was $t_s=3$ min and $TS=120$ min for NOC batches, giving as result 41 observations per batch. In addition, time was included as variable. The total batches (40) were variable-wise unfolded, centered and scaled, and joined in the training set.

A feature extraction step was applied by applying Multiway Principal Component Analysis (MPCA) to the centered and scaled NOC batches, keeping the same number of components than variables (9) and giving as result the projection model. The projection of both NOC and Abnormal Operating Conditions (AOC) batches onto such model produces the scores, used as input of both binary and multi-class classifiers. SVM are applied as classification algorithm.

The training sets are used as validation sets in both classifiers for adjusting parameters. In this way, the polynomial kernel of third order showed the best performance for the binary classifier and a polynomial kernel of fourth order for the multi-class classifier. On the other hand, a test set containing one batch per class is constructed and used for evaluating the performance of both classifiers. Table 3 and 4 show the corresponding results.

Table 3. Performance of the binary classification stage

Class	F1 score (%)	
	Validation set	Test set
Nominal	100	100
Faulty	100	100

Table 4. Performance of the multi-class classification stage

Class	F1 score (%)	
	Validation set	Test set
Nominal	100	100
Fault 1	100	100
Fault 2	100	100
Fault 3	100	100
Mean	100	100

Table 5. ND approach validation on two novel faults. Performance in terms of F1(%)

Novel fault	Detection stage		Diagnosis stage			
	NOC	Fault	NOC	Fault 1	Fault 2	Fault 3
Fault 4	-	100	100	-	-	-
Fault 5	-	98.8	7.3	-	-	92.7

The results in Tables 3 and 4 show the good performance of both detection and diagnosis stages of the ND approach. In order to assess this system, two batches with two novel faults are tested. These novel faults are:

- $C_0^{Fe(II)}$ higher (40ppm) than the nominal (10ppm) as fault 4

- Lack of irradiation source (configuration A) as fault 5.

The results after applying simultaneously both detection and diagnosis stages are reported in Table 5. As observed, the whole batch with fault 4 is diagnosed as a novel fault, indicated by its detection as fault by the detection stage and the lack of positive diagnosis of all the known faults (faults 1 to 3). On the contrary, fault 5 is detected as such but only few observations are diagnosed as novel fault and most of them diagnosed as fault three. This issue points out that the behavior of the on-line variables when there is no irradiation source is similar to their behavior when the H₂O₂ dosage protocol is different to the protocols included in the experimental design.

In general, successful results have been accomplished by applying the proposed ND approach to a Photo-Fenton process in pilot scale.

4. Conclusions

Novelty Detection approach has been proposed based on the combination of two classification stages: a binary or detection stage and a multi-class or diagnosis stage. The proposed approach has been applied to a Photo-Fenton process in a pilot plant where the contaminant to treat and degrade is coffee. Successful results have been obtained by detecting two novel faults not included in the constructed models. These results encourage to go on with further research and to use batches with different contaminants as novel faults. It would be probably required including the off-line measured variables in the training data so as to perform potential models that take into account the degradation behavior of the different contaminants. Such models might be then capable of detecting novel contaminants in wastewater treatments, which would constitute a powerful tool not only in FD area, but also in water research.

Acknowledgements

Authors would like to acknowledge Ministerio de Ciencia e Innovación and the European Regional Development Found for supporting the present research through project DPI2009-09386. We also appreciate financial support from Generalitat de Catalunya (FI program, 2011F1_B200181) and from Universidad de Carabobo (professorial grant CD-4352).

References

- K. Burbeck, S. Nadjm-Tehrani, 2007. Adaptive real-time anomaly detection with incremental clustering. Information Security Technical Report, 12:56-67.
- S. Ha An, G. Heo, S. Heung Chang, 2011. Detection of process anomalies using an improved statistical learning framework. Expert Systems with Applications, 38:1356-1363.
- A. Patcha, J-M. Park, 2007. An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer Networks, 51:3448-3470.
- J.J. Pignatello, E. Oliveros, A.A MacKay, 2006. Advanced oxidation processes for organic contaminant destruction based on the Fenton reaction and related chemistry. Critical Reviews in Environmental Science and Technology, 36:1-84.
- M. Tokumura, H. Tawfeek, Y. Kawase, 2008. Decolorization of dark brown colored coffee effluent by solar photo-Fenton reaction: Effect of solar light dose on decolorization kinetics. Water Research, 42:1665-4673.
- I. Gulkaya, A. Gulerman, F. Dilek, 2006. Importance of H₂O₂/Fe²⁺ ratio in Fenton's treatment of a carpet dyeing wastewater. Journal of Hazardous Materials, B136:763-769.
- C.J. Van Rijsbergen, 1979. Information retrieval (2nd ed.). London: Butterworths