

FAULT DIAGNOSIS UNDER MULTIPLE SEQUENTIAL FAULTS OF THE RAIN-GAUGE NETWORK USED TO CONTROL THE BARCELONA SEWER SYSTEM

Jaume Figueras, Vicenç Puig, Joseba Quevedo

*Automatic Control Department (ESAI) - Campus de Terrassa
Universitat Politècnica de Catalunya (UPC)
Rambla Sant Nebridi, 10. 08222 Terrassa (Spain)
jaume.figueras@upc.es*

Abstract: This paper discusses the problem of fault diagnosis under multiple sequential faults occurrence. In industrial applications this type of fault is the most common since the continuous operation of systems/processes is required. The fault diagnosis algorithms should cope with such type of multiple faults, but degradation in their fault isolation capabilities is introduced until the point that they should be stopped. A new algorithm to design the fault diagnosis system to be able to tolerate multiple sequential sensor faults is proposed. Finally, an example based on the rain-gauge sensor network of the Barcelona sewer system will be used to illustrate how the associated fault diagnosis system behaves under multiple sequential faults occurrence and to test the proposed algorithm. *Copyright © 2005 IFAC.*

Keywords: Sensor network analysis and design, Fault Diagnosis, Multiple Faults, Reconfiguration, Redundancy

1. INTRODUCTION

As noticed by Koscielny (2003), the on-line diagnosis of industrial processes is mostly performed by assumption of single faults. This significantly simplifies fault isolation algorithms. However, in large-scale technological installations with thousands of components such assumption is questionable. When with dealing multiple faults two types of situations are considered either if they appear as simultaneous or sequential. In this paper, the case of sequential faults is considered.

A fault diagnosis system is affected by multiple sequential faults since after the occurrence of fault has been diagnosed, the component affected by this fault is not available anymore. This is especially true in case of sensor faults. Since from structural analysis (Staroswiecki, 1989; 2000), it is known that redundancy relations are determined combining model equations with known variables, when a sensor/actuator is in fault, the set of redundancy relations will change.

This paper deals with static linear fault detection models derived from correlation analysis with the case study of the Barcelona rain-gauge network used for the global optimal control of retention tanks and the sewer system (Cembrano, 2004).

2. REVIEWING THE FAULT DIAGNOSIS PROCEDURE

2.1 FDI fault isolation

Model based fault detection tests are based on the evaluation of a set of relations $r_i(k)$, also known as **analytical redundancy relations** (ARRs), derived from the elementary models of system components and the available measurements coming from sensors available:

$$\mathbf{r}(k) = f(\mathbf{y}_{measured}(k), \mathbf{y}_{model}(k)) \quad (1)$$

Each detection test (r_i) should be evaluated on-line in order to decide if it is or not violated at a given time instant (typically a binary codification is used: 0 indicate not violation and 1 violation):

$$s_i(k) = \begin{cases} 0 & \text{if } |r_i(k)| < \tau_i \quad (\text{no fault}) \\ 1 & \text{if } |r_i(k)| \geq \tau_i \quad (\text{fault}) \end{cases} \quad (2)$$

where τ_i is the threshold associated to the detection test $r_i(k)$. This test constitutes the **detection phase** of the diagnosis process.

Finally, the evaluation of each detection test will provide the **actual** (or **observed**) **fault signature** of the system: $s(k) = [s_1(k), \dots, s_n(k)]$ that it is set of symptoms that will be supplied to the fault isolation module in order to isolate the fault.

2.2 FDI fault isolation

Given a set of symptoms $s(k) = [s_1(k), \dots, s_n(k)]$, and a set of considered faults: f_1, f_2, \dots, f_m , the **theoretical fault signature matrix**, Σ , can be defined binary codifying the effect or not of a fault in every symptom. This matrix has as many rows as symptoms and as many columns as faults are considered. The element Σ_{ij} of this matrix is equal to 1 meaning that the j^{th} fault appears in the expression of the i^{th} symptom generator, otherwise is equal to 0. Then, fault isolation will consist in looking for the **theoretical fault signature** in the **fault signature matrix** that matches with the observed signature. Let us consider an example. Given a set of three sensors such that: $y_1 = r_1(y_2)$, $y_2 = r_2(y_3)$ and $y_3 = r_3(y_1)$, the fault signature matrix is

	f_{y1}	f_{y2}	f_{y3}
s_1	1	1	0
s_2	0	1	1
s_3	1	0	1

Table 1. Fault signature matrix

Then, the following logical tests will allow isolating the different faults without considering that noise or perturbations may cause detection errors:

$$\begin{aligned} f_{y1} &= s_1 \wedge s_2 \wedge s_3 \\ f_{y2} &= s_1 \wedge \bar{s}_2 \wedge s_3 \\ f_{y3} &= s_1 \wedge s_2 \wedge \bar{s}_3 \end{aligned} \quad (3)$$

2.3 Fault isolability and the impact of sequence of faults

According to Gertler (1998), a set of faults are **isolable** in the fault signature matrix if their respective columns are different. For example, in *Table 1*, f_{y1} , f_{y2} and f_{y3} are isolable. However, when a fault appears in a given sensor, the fault signature matrix should be modified eliminating that sensor

from all the redundancy relations by substituting it by its model. Preceding in such a way the fault signature matrix will be modified and its isolability properties will be modified. For example, after a fault in sensor f_{y1} has been detected and isolated, the fault signature matrix should be adapted eliminating sensor y_1 from this table, in this case replacing sensor y_1 in model of y_3 by its model:

$$y_3 = r_3(r_1(y_2)) \quad (4)$$

The fault signature matrix given in *Table 1* will be modified accordingly

	f_{y2}	f_{y3}
s_2	1	1
s_3	1	1

Table 2. Fault signature matrix

As it can be seen from this new table, now faults in sensors y_2 and y_3 cannot be isolated. This replacement is algebraically possible because of the static and linearity properties of the sensor models for fault detection but there can be numerical singularities where it is impossible to substitute the sensor in fault with its model.

3. MINIMAL AND REDUNDANT SETS OF SENSOR

Following Staroswiecki (2004), let I given a set of sensors available in the non-faulty system, and $J \subseteq I$ a given set configuration with a subset of sensors active (non-faulty). Considering P the property of having a set of redundancy relations, which associated fault signature matrix, allows isolating all the sensors, and then two following sets can be introduced:

- the set 2^{I+} is defined by the configurations that satisfies $P(J)=I$, i.e., the configuration J satisfies that generates a set of redundancy relations that allow to isolate all the sensors involved.
- the set 2^{I-} is defined by the configurations that satisfies $P(J)=0$, i.e., such property is not satisfied.

Then, the following definitions can be introduced:

Definition 1. $J \in 2^{I+}$ is a **minimal sensor set** for property P (MSS) defined above, if and only if $\forall K \subseteq J, K \in 2^{I+}$.

Definition 2. $J \in 2^{I+}$ is a **redundant sensor set** for property P (RSS) above, if and only if it is not minimal.

The number of minimal and redundant sensor sets for a given set of sensors I is denoted respectively by:

$|MSS(I)|$ and $|RSS(I)|$. Given a sensor configuration J , if $|MSS(J)|=I$, then if any sensor that belongs to this set is lost the resulting set of sensors will not satisfy the isolability property, i.e., is a **critical sensor**. Staroswiecki (2004) also defines an evaluation measurement of the fault tolerance capability as the **Redundancy Degrees**.

Definition 3: Weak Redundancy Degree associated with a property P of a functional z and $J \subseteq I$

$$WDR(z, J) = |J| - \min_{K \in MSS(J)} |K| \quad (5)$$

Where WRD is, in our case, the longest sequence of faults before the system loses the isolability property. To determine the maximal number of sensors that can be lost while keeping the property P of the system the Strong Redundancy Degree is defined.

Definition 4: Strong Redundancy Degree associated with a property P of a functional z and $J \subseteq I$

$$SRD(z, J) = |J| - \max_{J^* \in RSS(J)} |J \setminus J^*| - 1 \quad (6)$$

Where SRD is, in our case, the shortest sequence of faults before the system loses the isolability property.

4. DESIGN OF THE FAULT DIAGNOSIS SYSTEM TO MAXIMIZE MULTIPLE FAULT TOLERANCE

On previous studies on the design of a fault diagnosis system for a sensor network using models obtained by correlation analysis (Puig 2003), the associated fault signature matrix was designed to provide the best fault isolation properties in case of single faults but there was no aim to provide an acceptable sequential multiple fault recoverability. Since the models were derived from correlation analysis, a certain level of correlation LC is fixed to determine which sets of sensors can build the fault detection models for each sensor. Let $S_i^{LC} \subset I$ the sets of sensors that are correlated at a level LC for the sensor i . Let ns be the number of sensors used to build the fault detection models. ns should be determined carefully, because a low ns implies good fault tolerance since the fault signature matrix will have a low number of 1's, but the model accuracy will be poor. On the other hand, choosing a high value for ns will fill the signature matrix of 1's in a way such that the fault tolerance capabilities will decrease but the model accuracy will increase. Then, it is needed that $\forall i \in I: |S_i^{LC}| \geq ns$ in order to build the models. If this condition is not true for all the sensors, then, the either level of correlation or the number of sensors

required is too high. Once the sensor sets are defined, a backtracking algorithm is used to build the models and trying to satisfy at the same time that the resulting fault signature matrix is column-canonical (Gertler 1998). Column-canonical signature matrices force the sensors to appear in the same number of models (ARRs), and this prevents the existence of critical sensors after the appearance of a first fault, i.e. $\neg \exists J \subset I: |J| = |I| - 1 \wedge J \in 2^{I-}$. Let no_j be the number of occurrences of the sensor j in all the sensor sets, then if $\exists j \in I: no_j < |I|nr$, it can be assured that there are sensors not enough correlated with the others, and no procedure can be applied to improve the multiple fault tolerance properties of the signature matrix because of a bad sensor network design or the existence of singularities in the network. From the correlation analysis the sensor network can be redesigned to improve its fault tolerance capabilities. The proposed backtracking algorithm acts as follows:

- (1) Build the $S_i^{LC} \subset I$ sets from correlation analysis
- (2) Order the sets in ascending order by its cardinal
- (3) Repeat for each sensor
 - Take a new group of ns sensors from its set and test the number of occurrences of each sensor in the models until none of the sensor occurrences is greater than ns and none of the groups is repeated or no group can be found.
 - If a group of ns sensors to build the model is found move to the next sensor.
 - If no group of ns sensors can be found move to the previous sensor.
 - If the actual sensor is the first then no solution can be found.
 - If the actual sensor is the last of the list then a solution has been found.
- (4) For each sensor built its model.
- (5) Build the signature matrix.

Now with a column-canonical signature matrix, it can be assured that the system is detectable and isolable even after the occurrence of a first fault. Once a fault is identified two different approaches can be used, a new signature matrix is calculated following *Section 2.3* or a new set of models and signature matrix can be built using the algorithm described before. The second approach lasts until no solution can be found, and no column-canonical signature matrix can be built, so at this point the only approach that can be used is described in *Section 2.3* substituting the faulty sensor by its model. It is also possible to relax the signature matrix to a weakly isolating structure, although the loss of uniformity of the signature matrix it is also detectable and isolable.

5. APPLICATION EXAMPLE: RAIN GAUGE NETWORK OF BARCELONA SEWER NETWORK

In this paper, fault diagnosis of sequential faults in the rain gauge network of Barcelona's urban sewer system is presented (Puig, 2003).

5.1 Modelling rain-gauges

In order to detect faults in a given rain-gauge P_{ut} ("rain-gauge under test"), first a model that relates its measurements with previous ones of the same rain-gauge ("auto-regressive AR model"):

$$P_{ut}(k) = f(P_{ut}(k-1), P_{ut}(k-2), \dots, P_{ut}(k-n)) \quad (7)$$

or with other rain-gauges spatially correlated ("moving-average MA model"):

$$P_{ut}(k) = f(P_1(k), P_2(k), \dots, P_m(k)) \quad (8)$$

should be derived. When an AR model is used as a means of providing analytical redundancy, it is said that it exploits *temporal redundancy* in data while a MA model exploits *spatial redundancy*. In case of linearity, AR and MA models can be viewed special cases of a more general relation based on a linear dynamic ARMA model with multiple inputs $\mathbf{u}(k) = [P_1(k), \dots, P_n(k)]$ and one output $\mathbf{y}(k) = P_{ut}(k)$ in discrete-time, then the nominal input-output relationship, without faults, disturbances and noise can be represented by

$$\mathbf{y}(k) = \mathbf{M}(q)\mathbf{u}(k) \quad (9)$$

where: $\mathbf{M}(q)$ is the discrete transfer function obtained from the input-output relationship using classical transform operator (q -operator). This general form will exploit *spatial-temporal redundancy*. Due to rain is a uncorrelated stochastic process, AR models are not useful in rain-gauge fault detection. On the other hand, MA models derived from the correlation analysis between all the existent rain gauges in the telemetry system will allow deriving which are the most correlated rain gauges with a given rain gauge under test. The use of a MA model would not imply only to identify the most correlated rain gauges but also the delay among rain gauges. Usually, the delay among rain gauges is very variable being impossible to determine a fixed delay valid for all the scenarios. For this reason only static MA models would be considered for fault detection. In Fig 1, the results of performing correlation analysis, for a 5-year rain data register, taking as rain gauge under test P_1 is presented. From this figure, it can be observed that the most correlated rain gauges are in decreasing

order: $P_{18}, P_{21}, P_8, P_{15}, P_4, P_{22}$, etc. Once the most correlated rain gauges have been derived the following question to answer is how many rain gauge to consider in order to build a fault detection model for a given rain-gauge. To answer this question, there is a compromise between fault detection and fault isolation model properties. An increase in the number of rain-gauges used to model the rain rain-gauge under test improves the prediction provided by the model, but decreases the capacity to isolate the faulty rain-gauge because the model is affected by the faults of more rain-gauges (Fig. 2). After some experimentation, three rain gauges, providing the 70% of data variance explained (Fig. 2), was the number selected in this application.

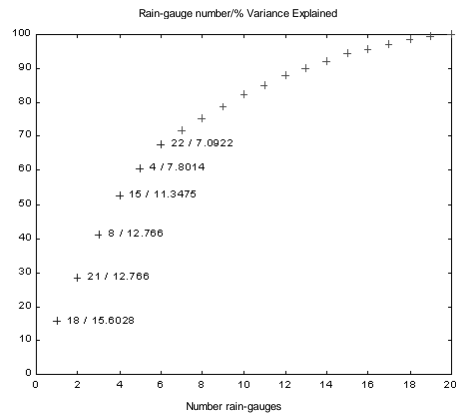


Fig. 1. Rain gauges most correlated with rain gauge P_1

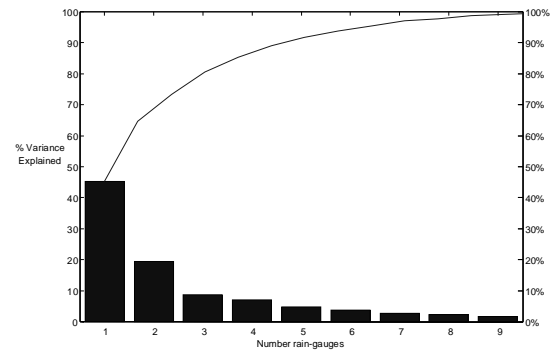


Fig. 2. % of data variance explained with respect to the number of rain-gauges for a given rain-gauge using Principal Component Analysis (PCA)

In the case of Barcelona's rain gauges the fault detection models for every rain gauge under test have the Boolean signature presented in Fig. 3. With no other aim than fault isolability, signature matrix of Figure 3 has two critical sensors, P_{13} and P_{19} , that in case of fault will derive in a sensor network does not satisfy such property. The, as defined by Staroswiecki (2004), the Strong Redundancy Degree (SRD_M) for the signature matrix M of the system is 1, so there is the possibility of system failure after the fault of one single sensor.

The probability of system failure (FP) is defined as the relation between the number of sensor fault combinations that lead to a system failure and the total number of sensor combinations. Using this probability, it can be determined which fault signature matrix is more fault tolerant reliable regarding the fault isolability property. After a sensor fault, a new matrix has to be calculated as described in Section 2.3. So, for a n sensor network there is a set NM , with $|NM| = n$ of possible new matrices.

	P1	P2	P3	P4	P5	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	
P1	X						X										X					
P2		X																				
P3			X																			
P4				X																		
P5					X																	
P6						X																
P7							X															
P8	X							X														X
P9									X								X	X	X			
P10						X				X												
P11											X											
P12		X										X										
P13			X										X									
P14				X										X								
P15		X													X						X	X
P16			X													X						
P17				X													X					
P18	X																	X				X
P19			X	X															X			
P20				X	X															X		
P21					X																X	
P22		X																				X

Fig. 3. Fault signature matrix of the rain-gauge network of Barcelona's sewer system

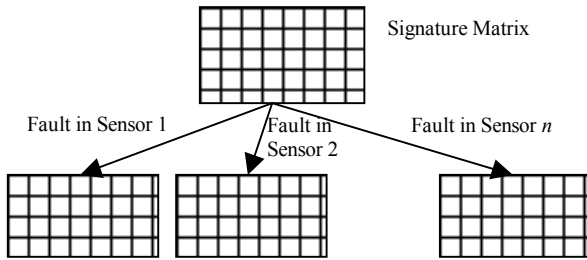


Fig. 4. New Matrices after a sensor fault

The probability is computed as the number of matrices that are isolable with respect of the number of new possible matrices.

$$FP(M) = 1 - \frac{|ISOLABLE(NM)|}{|NM|}$$

$FP(M)=0$ means that all possible new matrices are isolable, $FP(M)=0.2$ mean that a 20% of new possible matrices are not isolable after a sensor fault. Following this definition the original signature matrix (Figure 3) has $FP(M)=0.0909$ after one sensor fault. Figure 5 is an extract of the correlation table of a statistical study of all the rain gauge in the Barcelona network. From this table and the algorithm described in Section 4, a set of new signature matrices (NSM) can be build in order to study how fault tolerant are, how many sensors should be used to build the detection models, and how good are these detection models. After analyzing thousands of signature matrices, the matrices build with 2 sensors in the fault detection model are more fault tolerant than matrices build with 4 sensors in the fault

detection model. 82% of new fault signature matrices with 2 sensors will not fail after one sensor fault, 61% of new signature matrices built with 3 sensors will not fail after one sensor fault while all new matrices built with 4 sensors lead to a system failure after one sensor fault.

	1	%	2	%	3	%	4	%
P01	18	15.60	21	12.76	8	12.76	15	11.34
P02	15	17.02	12	15.60	13	14.18	22	12.05
P03	19	26.08	16	16.66	17	7.97	4	7.24
P04	21	22.22	20	17.77	19	9.62	3	7.40
P05	20	17.24	14	14.94	13	14.94	12	8.04
P06	11	26.66	14	20.74	10	17.77		
P07	3	14.39	10	11.36	19	11.36	11	9.84
P08	1	21.01	15	13.04	22	12.31	18	10.86
P09	17	20.13	19	15.97	18	15.97	3	6.94
P10	11	25.18	6	22.22	14	12.59	16	5.92
P11	6	28.88	10	18.51	14	14.07	16	6.66
P12	2	23.07	13	22.22	15	11.96	22	9.40
P13	2	14.89	12	13.47	5	11.34	20	11.34
P14	6	21.01	16	13.76	11	12.31	20	7.97
P15	2	18.75	22	18.05	21	12.50	1	9.02
P16	3	15.21	6	13.04	19	10.86	11	10.86
P17	9	22.91	19	19.44	3	11.80	18	6.94
P18	1	15.27	21	14.58	9	13.19	19	7.63
P19	3	23.91	17	15.21	4	11.59	16	7.97
P20	4	18.84	5	12.31	13	10.86	21	7.97
P21	4	19.14	15	13.47	18	11.34	20	9.92
P22	15	24.82	2	16.31	21	8.51	12	7.80

Fig. 5 Extract from the correlation table of the Barcelona rain gauge network

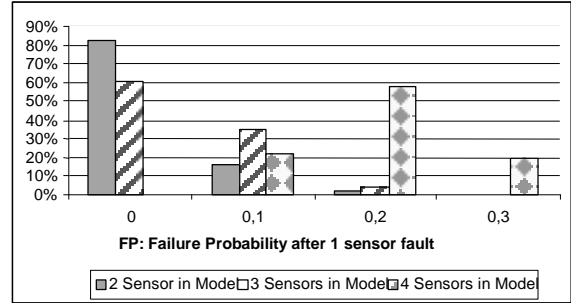


Fig. 6. System failure probability after one sensor fault for different types of signature matrices.

On the other hand, model quality can be computed by the sum of correlation values of the sensors used to build the models. Calculating the average with respect of the number of sensors lead us to the mean model quality (MMQ)

$$MMQ = \frac{\sum_{i=1}^n \sum_{j=1}^k Corri^j}{n}$$

Where n is the number of sensors of the system, k is the number of sensors in the model and $Corri^j$ is the value of correlation of the sensor i with j . From this definition the signature matrix proposed in Figure 3 has a MMQ of 16.26. The MMQ for the matrices build with 3 sensors in the fault detection model have a higher MMQ than the others, FP has a nearly inappreciable influence on MMQ , although a lower FP means low MMQ and higher FP means high MMQ for matrices build with 3 sensors. From these results a new set of models and a new signature matrix shown in Figure 8 was calculated and two sensors were eliminated due to a lack of relevance. Studying the probability of system failure after consecutive sensor faults (Table 2) using the new fault signature matrix of Figure 8 and the original

matrix of Figure 3, the improvement of the system fault tolerance capabilities is significant just choosing the right models for fault detection.

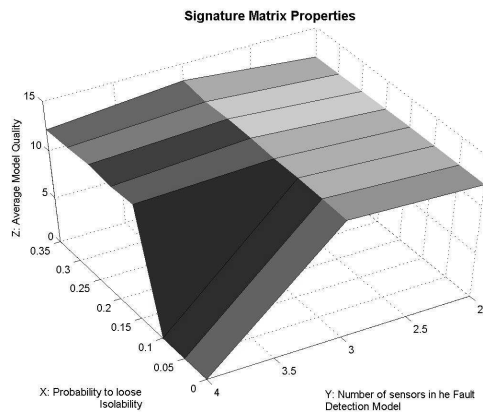


Fig. 7. Quality versus number of sensors and failure probability.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20
P1	X																			
P2		X								X	X									X
P3			X																	
P4				X									X		X					
P5					X					X										
P6						X				X	X									
P7							X									X				
P8								X	X	X			X							
P9									X	X										
P10		X								X	X									X
P11					X					X	X									
P12						X			X		X	X							X	
P13		X								X			X							X
P14			X										X							X
P15								X		X				X		X	X			
P16	X			X											X	X	X			
P17				X										X	X	X	X			
P18					X													X	X	
P19	X																X	X	X	
P20		X											X					X	X	X

Fig. 8. Fault signature matrix of the rain-gauge network of Barcelona’s sewer system applying fault tolerance techniques

	Old Matrix	New Matrix
1 Consecutive Fault	9,09%	0,00%
2 Consecutive Faults	18,18%	1,58%
3 Consecutive Faults	26,96%	5,64%
4 Consecutive Faults	35,24%	12,77%
5 Consecutive Faults	43,00%	23,12%

Table 2. Comparison of the probability of system failure after consecutive faults

6. CONCLUSIONS AND FURTHER WORK

In this paper, the problem of reconfiguring the fault diagnosis system in order to cope with multiple sequential faults is addressed. After a fault occurs, the fault signature matrix should be modified eliminating the faulty sensor and replacing its occurrences in the rest of equations by its model. Doing such a process fault isolability properties of the fault signature matrix will be modified. Using the methodology proposed by Staroswiecki (2004), the

reconfiguration of the fault diagnosis systems of the rain-gauge network of Barcelona’s sewer system is analysed. A new algorithm to design the fault diagnosis system taking into account the tolerance to sequential faults is presented and test on such sensor network. As a further research, new policies of model reconfiguration after a sensor fault, and searching a new signature matrix are under study. The difficulty in this approach is in the number of correlated sensors, since the level of correlation is fixed, the number of correlated sensors is also fixed, and in many cases there are not enough correlated sensors to build the models. Finally, the algorithms used to calculate and evaluate the new models and matrices are very time consuming because of combinatorial explosion, although computers are faster every day and system design is an off-line work, a more optimized algorithms are tested in order to reduce computer time.

ACKNOWLEDGMENTS

The authors wish to thank the support received by the Research Commission of the Generalitat of Catalunya (ref. 2001SGR00236) and by Spanish CICYT (ref. DPI2002-02147) and the support from CLABSA, the company in charge of the sewer system management in Barcelona, for supplying and validating all data we needed for the optimal control of the sewage system of Barcelona.

REFERENCES

- Cembrano, G. et al. (2004). “Optimal control of urban drainage systems. A case study”. *Control Engineering Practice*, 12 (1), pp. 1-9.
- Gertler J.J. (1998). “Fault Detection and Diagnosis in Engineering Systems”. Marcel Dekker.
- Koscielny, J.M., Bartys, M. “Multiple Fault Isolation in Diagnostics of Industrial Processes”. *European Control Conferece (ECC03)*. Cambridge, UK.
- Puig V., Quevedo J., Figueras J., Riera S., Cembrano G., Salamero M., Wilhelmi G. (2003). *Fault detection and isolation of rain gauges and limnimeters of Barcelona's Sewer system using interval models IFAC SAFEPROCESS, 2003* Washington, USA.
- Staroswiecki, M., P. Declerck. (1989) “Analytical redundancy in non-linear interconnected systems by means of structural analysis”. *IFAC Advanced Information Processing in Automatic Control (AIPAC'89)*, pp. 51-55. Nancy. France.
- Staroswiecki, M., J. P. Cassar, P. Declerk (2000). “A structural framework for the design of FDI system in large scale industrial plants”. In “*Issues of Fault Diagnosis for Dynamic Systems*”. Patton, R.J., Frank, P.M., Clark, R.N. (eds). Springer-Verlag.
- Staroswiecki, M., Hoblos, G., Aitouche, A. (2004). “Sensor network design for fault tolerant estimation”. *International Journal of Adaptive Control and Signal Processing*, 2004.