SENSOR FAULT DETECTION IN A REAL HYDRAULIC SYSTEM USING A CLASSIFICATION APPROACH

Oriane Le Pocher, Eric Duviella
Univ. Lille Nord de France, F-59000 Lille, France
EMDouai, IA, F-59500 Douai, France

Karine Chuquet
VNF - Service de la navigation du Nord Pas-de-Calais, 37 rue du Plat, 59034 Lille Cedex, France

Keywords: Supervision, Fault detection, Classification algorithm, Large scale system, Hydraulic system.

Abstract: This paper focuses on the sensor fault detection of a hydraulic channel used for navigation. This system has the particularities to have large scale dimension, without slope, with several inputs and ouputs, and thus difficult to be modelled according to classical modelling methods. For recent years, it was equipped with level sensors in order to have better knowwledge of its behavior, to detect its state online and thus improve its management. However, level sensors are subjected to measurement or transmission errors, setting errors, and quick or slow drifts. In order to detect these sensor errors, a classification approach is proposed. It appears adapted to the fault detection of large scale hydraulic systems without model. The classification approach is used on data measured from 2006 to 2009. The first results and analysis show that the classification method is effective for addressing the problem of sensor fault detection.

1 INTRODUCTION

Hydrographical networks are large scale systems characterized by nonlinear dynamics and varying time delays. They are used for several human activities, especially navigation and transport. In the northern Europe, navigation channels assure the transport of goods with the objective, within a few years, of accomodating large broad gauge boats. The control of the water level in navigation channels becomes crucial. In order to achieve this objective, sensor networks have been implemented. These sensors allow the measurement of water levels or water discharges, and the implementation of level control algorithms for a local water management. At a larger scale, the level and discharge measurements are essential to provide an efficient water management of the navigation channel networks, by mainly characterizing their state online. However, sensor networks are impacted by measure errors, transmission faults, or drifts of operation. So, in order to improve the management of navigation channels, sensor fault detection techniques have to be employed.

Fault Detection and Isolation (FDI) techniques are largely proposed in the litterature and employed by following a systematic approach. The first step consists in characterizing the operating modes of the system to be supervised. Several model-based approaches were proposed (Frank et al., 2000), based on parameters identification technique (Weihua et al., 2003), parity equations method (Gertler, 1998), diagnosis observers (Akhenak et al., 2004), or Kalman filters (Xie et al., 1994). Even if these FDI techniques have proven to be as powerful and effective, they require an accurate model of the system by minimizing the uncertainties and the process noise. Very recently, fault detection methods based on residual generation, extended Kalman filter and finite memory observer are proposed in (Bedjaoui and Weyer, 2010), in order to detect and localize leak in an irrigation network. This detection method is based on physical hydraulic system model, in particular on the Saint-Venant partial differential equations (Chow et al., 1998). However, due to their physical characteristics, *i.e.* large dimensions, no slope, etc., navigation channels can not always be modelled using physical laws without requiring numeric models. In this way, traditionnal FDI techniques cannot reach the fault detection aims.

When the physical modelling of the system is not realizable, pattern recognition techniques consti-

Le Pocher O., Duviella E. and Chuquet K..
SENSOR FAULT DETECTION IN A REAL HYDRAULIC SYSTEM USING A CLASSIFICATION APPROACH.
DOI: 10.5220/0003571603820387
In *Proceedings of the 8th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2011), pages 382-387
ISBN: 978-989-8425-74-4

tute an interesting alternative approach for fault detection problem. They consist in extracting information on the system state by using the signals collected from sensors (Hartert et al., 2010). The operating modes of the system are represented by classes which are built according to dynamic classification algorithms, as the CDL algorithm (Cluster Detection and Labeling), algorithms based on adaptive resonance theory (ART) networks (Su and Liu, 2005), (Eltoft and de Figueiredo, 1998), FMC algorithm (Fuzzy Min-max Clustering) (Mouchaweh et al., 2002) or AUDyC (Auto-Adaptive Dynamical Clustering) algorithm (Lecoeuche et al., 2004). In (Traore et al., 2009), the AUDyC algorithm is employed to supervise a thermo-regulator system subjected to slow and quick drifts in its dynamics.

This paper focuses on sensor fault detection of the Cuinchy-Fontinettes navigation channel. This system, located in the north of France, is characterized by large dimensions and no slope. Thus, it is not possible to be modelled according to physical approach. Thus, a classification approach is proposed to address the fault detection problem. In section II, the real hydraulic system and the problem of sensor fault detection are presented. The classification approach based on AUDyC algorithm is detailed in section III. This fault detection technique allows the determination of indicators characterizing the real-time drift of sensors. Finally, in section V, the proposed approach is applied on real data measured from 2006 to 2009, and its performance for the detection of sensor faults are highlighted.

2 PROBLEM STATEMENT

The channel studied is the reach Cuinchy-Fontinettes which is located in the North of France between the lock of Cuinchy at the East of the town Bethune and, at the Southwest of the town Saint-Omer, the lock of Fontinettes (*see* Figure 1). With 42.3 *km* long and 51.8 *m* large, this reach is part of the broad gauge river network of North of France. It is characterized by no significant slope. It can handle boats until 185 *m* long and 11.4 *m* large. This channel is entirely artificial. The first part of the channel, *i.e.* 28.7 *km* from Cuinchy to Aire sur la Lys, is called ”canal d’Aire” and has been built in 1820. The second part of the channel, *i.e.* 13.6 *km* from Aire sur la Lys to Saint Omer, is called ”canal de Neuffossé” and has been built in the eleventh century.

The Cuinchy-Fontinettes is managed by VNF (Voies Navigables de France). The role of VNF is to maintain the level of the channel at $NNL = 19.52$ *m* to allow the navigation (NNL = Normal Navigation Level). The main issue for the management of this reach is to counterbalance the navigation flow of 3 m^3/s at Fontinettes. A part of the water comes from the navigation flow at Cuinchy which represents 0.6 m^3/s. The difference between the two navigation flows comes from the size of the lock at Fontinettes which is 13 *m* high whereas the lock at Cuinchy is only 2 *m* high. The counterpart of water comes from different hydrants and discharges. A high number of rivers are inverted siphon and pass under the reach but three of them feed directly the reach. Another solution to feed the reach is the Cuinchy gate. This gate allows a controlled feeding of the reach with the water of the Deûle river. In the same way, the gate called ”Porte de Garde” at Aire sur la Lys allows controlled exchanges between the Lys river and the reach. Finally a high number of anthropogenic discharges (more than 320) feed the reach in an unknown way.

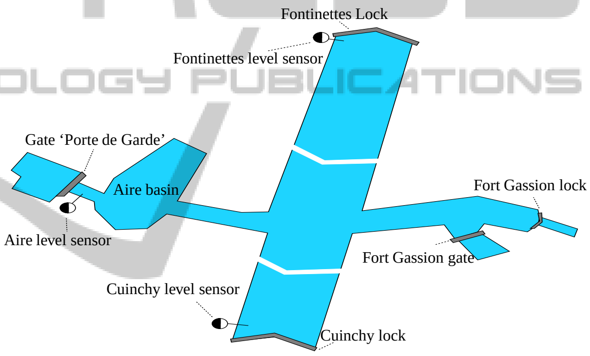

Figure 1: Scheme of the navigation canal Cuinchy-Fontinettes.

For recent years the reach has been equipped with level sensors in order to better know its behaviour. The level sensors used for this study are located downstream from the lock of Cuinchy, upstream from the lock of Fontinettes and in the Aire basin. The Cuinchy and Aire sensors are composed of an ultrasonic transducer linked with a level transformer. The Fontinettes sensor is a Probe with a transducer integrated. The technology used is based on ultrasonic sound and allows the processing of echoes. The data processing and transmission are realized by teleprocessing equipments. The three sensors deliver the mean of the levels measured every quarter hour. The waves due to navigation or Fontinettes lock operations are averaged.

Sensors are subjected to the weather and the environment and as every electronic device can break down or be impacted over time. Several types of errors can occur. A bad setting of the sensors can lead to systematic errors. Aberrant data are caused by local and temporal errors. Blockade of data can be due to a transmission fault for example. Level sensor can be subjected to slow temporal drifts. In order to use reliable measured data, it is essential to propose a fault detection technique.

The main problem for the fault detection technique proposal is the major difficulty of modeling the Cuinchy-Fontinettes channel without numeric approach. A fault detection technique by a pattern recognition approach is proposed in the next section in order to be freed from a model of the channel.

3 FAULT DETECTION BY A PATTERN RECOGNITION APPROACH

The fault detection method proposed in this paper is based on a classification technique. This classification technique consists in characterizing an operating mode of the dynamic system by a Gaussian model which constitutes a class. A class is determined according to pertinent selected data which present same similarities. According to these selected data, a Representation Space can be built, and the class can be represented in this space. Thus, the class of the normal operating mode, denoted C_n, can be determined (*see* Figure 2). A new class is create when a sufficient number of points is present in an area of the Representation Space. The new class, which is updated or created, is denoted evolutionary class C_e. It corresponds to a new operating mode.

When a measurement or transmission error occurs, a new point appears in the Representation Space far from the normal class C_n (*see* Figure 2.a). This point has to be detected and rejected if it is isolated. When the level sensor is subjected to slow drifts, the class updates online (*see* Figure 2.b). The characteristics of the normal class evolve during time. Finally, catalectic failures lead to a jump in the representation space (*see* Figure 2.c).

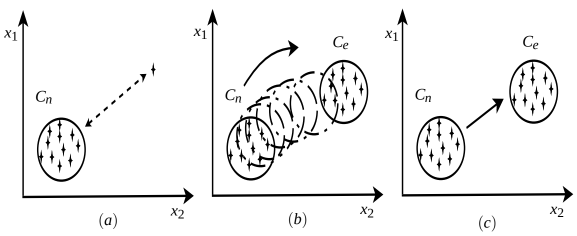

Figure 2: (a) Measurement or transmission errors, (b) slow drifts, (c) jump characterising catalectic failure, in a two dimensions representation space.

The classification technique which is proposed to monitor slow drifts and jumps, is based on the AUDyC algorithm. AUDyC is an evolutionary data classification algorithm whose role is to model, in a continuous way, the operating modes of dynamical systems. The technique is inspired from the mixed Gaussian model (Lecoeuche et al., 2004). The Gaussian classes are represented by prototypes P^j characterized by a center and a matrix of covariance. The prototypes characteristics are adapted to each new observation $X = (x_1, x_2, \dots x_n)$, with n the number of pertinent data, by using rules of recursive update on a sliding window of size N_{fen}, and by considering the totality of the prototype P^j, noted Card (P^j) on previous instant $k-1$, according to the algorithm described below. A new observation X is rejected if is too far from the current class. In other case, this observation is assigned to one of the N existing classes according to an adaptation procedure of the prototypes.

- If $Card(P^j)$=nb $< N_{fen}$: add information

$$\begin{cases} M_{P^j}(k) = M_{P^j}(k-1) + \dfrac{1}{nb+1}(X_k - M_{P^j}(k-1)), \\ \Omega_{P^j}(k) = \dfrac{nb-1}{nb}\Omega_{P^j}(k-1) + \\ \qquad \dfrac{1}{nb+1}(X_k - M_{P^j}(k-1))^\top (X_k - M_{P^j}(k-1)). \end{cases} \tag{1}$$

- If nb $\geq N_{fen}$: add or retreive information

$$\begin{cases} M_{P^j}(k) = M_{P^j}(k-1) + \dfrac{1}{N_{fen}}(\delta X^+ - \delta X^-), \\ \Omega_{P^j}(k) = \Omega_{P^j}(k-1) + \\ \qquad \Delta X \begin{bmatrix} \frac{1}{N_{fen}} & \frac{1}{N_{fen}(N_{fen}-1)} \\ \frac{1}{N_{fen}(N_{fen}-1)} & -\frac{(N_{fen}+1)}{N_{fen}(N_{fen}-1)} \end{bmatrix} \Delta X^\top, \end{cases} \tag{2}$$

where

$$\begin{cases} \delta X^+ = X^{new} - M_{P^j}(k-1), \\ \delta X^- = X^{old} - M_{P^j}(k-1), \\ \Delta X = [\delta X^+ \quad \delta X^-], \end{cases} \tag{3}$$

with $M_{Pj}(k)$ and $\Omega_{Pj}(k)$ respectively the center and the covariance matrix of the prototype P^j at instant k, N_{fen} the width of the slipping window, X^{new} and X^{old} new and old observation vectors, respectively.

This fault detection method based on AUDyC algorithm allows rejecting measurement or transmission errors, following slow drifts and catalectic failures. In order to detect drifts, fault indicators have to be calculated. A first indicator corresponds to the mean of the measured levels on the sliding window N_{fen}. It allows the visualisation of quick and slows drifts. A second indicator consists in computing the distance between the evolutive class C_e and the normal class C_n. The normal class is taken as reference. A fault occurence can lead to the drift of the evolutive class C_e from the normal class C_n. Increasing of this distance reveals, in some cases, the presence of faults. Amongst the several existing metrics, the Euclidian distance, denoted $d(M_e, M_n)$, is considered:

$$d(M_e, M_n) = \sqrt{(M_e - M_n)(M_e - M_n)^\top}, \qquad (4)$$

with M_e is the center of the class C_e and M_n the center of the normal class C_n. The center of the normal class M_n is fixed.

Moreover, by considering, for causal systems, that the outputs depend on the variation of the inputs, correlation indicators between the center of the evolutive class according to each direction of the Representation Space, denoted $K_{i,j}$ for data x_i and x_j, can reveal a fault occurence. That leads to compute correlation coefficients between the mean of each measured level on a sliding window of size N_{fen}. In a second step, error indicators are computed between the center of the evolutive class amongst each direction of the Representation Space also. The error indicators are denoted $\varepsilon_{i,j}$ for data x_i and x_j. Finally, the quadratic error indicators $\varepsilon^2_{i,j}$ are computed according to $\varepsilon_{i,j}$.

Quadratic error indicators $\varepsilon^2_{i,j}$ are used in order to detect setting errors. Error indicators $\varepsilon_{i,j}$ are used in order to detect slow drifts and to determine which sensor is faulty. Finally, correlation indicators $K_{i,j}$ allow the detection of quick drifts and catalectic failures. The redondancy of the indicators and a cross-comparison lead to determine which of the sensors is faulty. The fault detection approach is applied in the case of the Cuinchy-Fontinettes Channel on data from 2006 to 2009. Results and analysis are presented in the next section.

4 FAULT DETECTION IN THE CUINCHY-FONTINETTES CHANNEL

Measured data on the Cuinchy-Fontinettes Channel correspond to x_a, x_c and x_f, for Aire in the middle of the channel, for Cuinchy at the upstream, and for Fontinettes at the downstream, respectively. These data are measured with a sample time equal to 15 minutes from 2006 to 2009. It represents more than 135000×3 values. These data were not recorded at the same time. Indeed, there are discrepancies of few minutes between measurements. Then, the first step is to resynchronize all the data.

The second step consisted in building the Representation Space with the three measured data x_a, x_c and x_f, and to represent the normal class C_n (*see* Figure 3). The normal class is built with accurate measured data from April to June 2006 according to a sliding window N_{fen} equal to 2500 values. The center of the class C_n is around zero (relative levels according to the NNL), *i.e.* $[-0.034\ -0.058\ -0.059]$, and its covariance matrix is equal to:

$$\Sigma_n = \begin{bmatrix} 0.0060\ 0.0052\ 0.0055 \\ 0.0052\ 0.0055\ 0.0052 \\ 0.0055\ 0.0052\ 0.0065 \end{bmatrix}. \qquad (5)$$

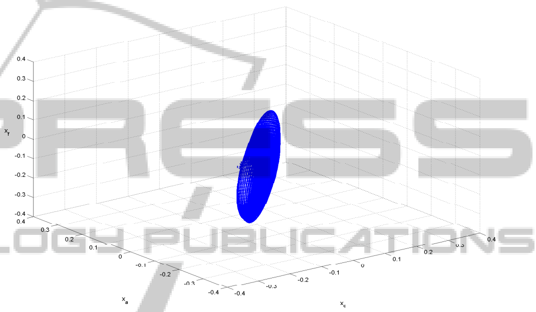

Figure 3: Normal class C_n in the Representation Space.

Although the total measured data was used, only years 2006 and 2009 were shown to highlight the performance of the proposed approach. Figure 4 shows the measured levels, *i.e.* x_c, x_a and x_f, and the detection and isolation of wrong measured data during the year 2009. The wrong data are depicted by black cross in Figure 4. The classification approach allows to reject these points automatically.

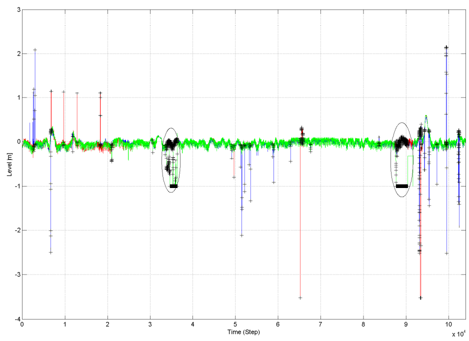

Figure 4: Data x_c (continuous line), x_a (dotted line) and x_f (dashed-dotted line) measured on 2009, and isolate wrong data (black cross).

Figure 5.*a* shows the measured levels in Cuinchy (blue continuous line), in Aire (red dashed line) and in Fontinettes (green dashed-dotted line) during 2009. The distance $d(M_e, M_n)$ is depicted in Figure

5.*b*. The correlation indicators $K_{c,a}$, $K_{c,f}$ and $K_{a,f}$, the quadratic error indicators $\varepsilon^2_{c,a}$, $\varepsilon^2_{c,f}$ and $\varepsilon^2_{a,f}$, and the error indicators $\varepsilon_{c,a}$, $\varepsilon_{c,f}$ and $\varepsilon_{a,f}$, are computed between the mean of measured levels in Cuinchy and Aire, in Cuinchy and Fontinettes, in Aire and Fontinettes, respectively. Indicators between Cuinchy and Aire are depicted in blue continuous line, those between Cuinchy and Fontinettes in red dashed line, and those between Aire and Fontinettes in green dashed-dotted line, in Figure 5.*c*, 5.*d* and 5.*e*, respectively. A threshold defined equal to 0.7 is considered in order to detect fault when one of the correlation indicators is under this threshold (*see* Figure 5.*c*).

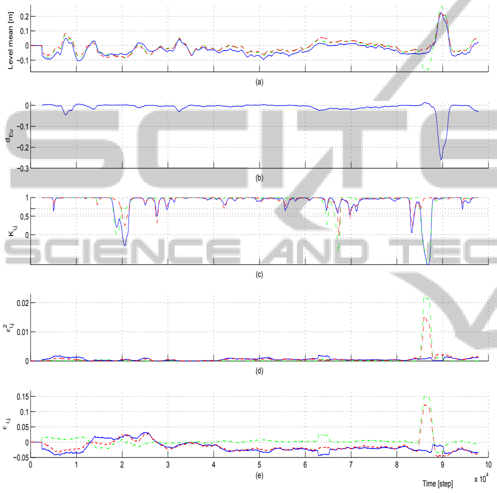

Figure 5: (a) Measured levels x_c (continuous line), x_a (dotted line) and x_f (dashed-dotted line), (b) Euclidienne distance $d_{Eu}(M_e, M_n)$, (c) correlation indicators $K_{c,a}$ (continuous line), $K_{c,f}$ (dashed line) and $K_{a,f}$ (dashed-dotted line), (d) quadratic error indicators $\varepsilon^2_{c,a}$ (continuous line), $\varepsilon^2_{c,f}$ (dashed line) and $\varepsilon^2_{a,f}$ (dashed-dotted line), (e) error indicators $\varepsilon_{c,a}$ (continuous line), $\varepsilon_{c,f}$ (dashed line) and $\varepsilon_{a,f}$ (dashed-dotted line), measured on 2009.

During 2009, by considering the distance $d(M_e, M_n)$ (*see* Figure 5.*b*), only one period around the 90000th step, is relevant to significant drift of the class C_e. However, around the 90000th sample, others indicators are close to their objective values. This means that there is no fault. In this period, the distance $d(M_e, M_n)$ is increasing because there is a modification of the operating mode, which can be the consequence of flood (*see* measured levels in Figure 5.*a*). Figure 5.*c* shows three periods where correlation indicators are under the fixed threashold, *i.e.* around 20000th, 66000th and 88000th samples. For the two first periods, there are no significant errors ε^2 and ε (*see* Figure 5.*d*, 5.*e*). During the third period around the 88000th sample, there are significant errors $\varepsilon^2_{c,f}$, $\varepsilon^2_{a,f}$, $\varepsilon_{c,f}$ and $\varepsilon_{a,f}$. Errors $\varepsilon^2_{c,a}$,and $\varepsilon_{c,a}$ are close to their objectives. This is relevant of slow drift on the level sensor in Fontinettes.

Figure 6 shows the measured levels, *i.e.* x_c, x_a and x_f, during the year 2006. The detected wrong measured data appear during the period around the 55000th sample (*see* black cross).

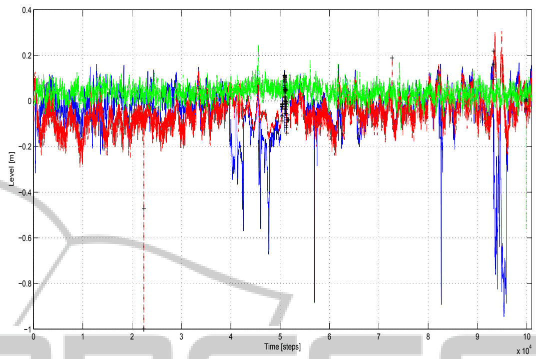

Figure 6: Data x_c (continuous line), x_a (dotted line) and x_f (dashed-dotted line) measured on 2006, and isolate wrong data (black cross).

The same indicators are determined for year 2006 and depicted in Figure 7, and the same conclusions can be obtained if the distance $d(M_e, M_n)$ and correlation indicators are taken into account. The most interesting point to show is the detection of slow drift of the Cuinchy level sensor during all the year. From the beginning of year 2006 to the 26000th sample, quadratic errors $\varepsilon^2_{c,a}$ and $\varepsilon^2_{c,f}$ are constant and around 0.0025 and $\varepsilon^2_{a,f}$ is close to zero (*see* Figure 7.*d*). It means that there is a setting error on the level sensor in Cuinchy. The setting error is evaluated from 0.05 m according to the errors $\varepsilon_{c,a}$ and $\varepsilon_{c,f}$ (*see* Figure 7.*e*). Then from the 26000th sample to the 65000th sample, all the errors are close to zero. It is possible to assume that the Cuinchy level sensor is correctly set. Finally, from the 65000th sample, quadratic errors $\varepsilon^2_{c,a}$ and $\varepsilon^2_{c,f}$ are increasing. It is relevant of slow drift of the Cuinchy level sensor. Errors $\varepsilon_{c,a}$ and $\varepsilon_{c,f}$ are decreasing to reach -0.1 m.

The fault detection method proposed in this article allows the detection of error setting, slow and quick drifts. It can be implemented online in order to detect these types of faults in real-time.

5 CONCLUSIONS

The sensor fault detection of real large scale systems without model is an interesting research problem. The well-known classical FDI techniques cannot be applied due to the difficulty of modelling. Thus, a fault detection approach without model is proposed in or-

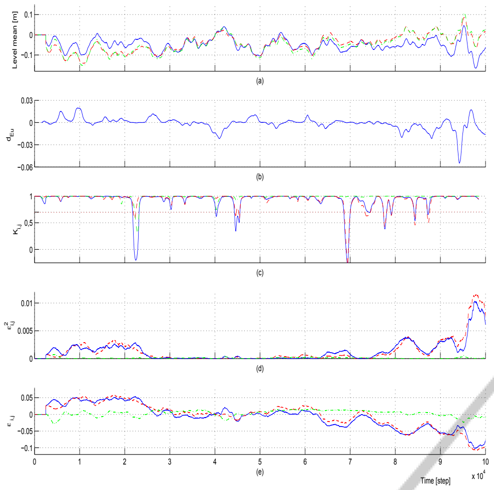

Figure 7: (a) Measured levels x_c (continuous line), x_a (dotted line) and x_f (dashed-dotted line), (b) Euclidienne distance $d_{Eu}(M_e, M_n)$, (c) correlation indicators $K_{c,a}$ (continuous line), $K_{c,f}$ (dashed line) and $K_{a,f}$ (dashed-dotted line), (d) quadratic error indicators $\varepsilon^2_{c,a}$ (continuous line), $\varepsilon^2_{c,f}$ (dashed line) and $\varepsilon^2_{a,f}$ (dashed-dotted line), (e) error indicators $\varepsilon_{c,a}$ (continuous line), $\varepsilon_{c,f}$ (dashed line) and $\varepsilon_{a,f}$ (dashed-dotted line), measured on 2006.

der to reach these objectives. The technique of supervision which is presented in this article is based on the Pattern recognition AUDyC algorithm. It has the advantage to limit physical knowledge of the system, and aims to modelling the operating modes of dynamical systems using only measured data. The characteristics of the operating mode are updated in real-time in order to follow the drifts due to sensor faults, and detect setting errors and measurement or transmission errors. The proposed technique is applied on a real hydrographical system with presents the particularities to not being modelled according to classical modelling methods. Fault indicators are determined according to levels which are measured since 2006. The first obtained results highlight the efficiency of the proposed fault detection method. However, these results have to be improved. The futur purposes consist in proposing more pertinent fault indicators by considering the measured upstream and downstream in the Cuinchy-Fontinettes channel. It should be also interesting to take into account the unknown inputs which correspond to overflows in the channel. In future works, a prognosis approach will be proposed to predict the future state of the level sensors in order to detect as soon as possible sensor faults. Finally, an implementation of the proposed technique on the real system may be considered at term.

REFERENCES

Akhenak, A., Chadli, M., Ragot, J., and Maquin, D. (2004). State estimation of uncertain multiple model with unknown inputs. In *43rd IEEE Conference on Decision and Control, Atlantis, Bahamas*, page 35633568.

Bedjaoui, N. and Weyer, E. (2010). Algorithms for leak detection, estimation, isolation and localization in open water channels. *Control Engineering Practice, In Press*.

Chow, V. T., Maidment, D. R., and Mays, L. W. (1998). *Applied Hydrology*. McGraw-Hill.

Eltoft, T. and de Figueiredo, R. (1998). A nez neural network for cluster-detection-and-labeling. *IEEE Trans. Neural Networks*, 9:1021–1035.

Frank, P. M., Ding, S. X., and KCipper-Seligcr, B. (2000). Current developments in the theory of fdi. In *SAFEPROCESS00, Budapest, Hungary*, pages 16–27.

Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Dekker.

Hartert, L., Mouchaweh, M. S., and Billaudel, P. (2010). *Intelligent Industrial Systems: Modeling, Automation and Adaptive Behavior*. IGI.

Lecoeuche, S., Lurette, C., and Lalot, S. (2004). New supervision architecture based on on-line modeling of non-stationary data. *Neural Computing and Applications Journal*, 13:323–338.

Mouchaweh, M. S., Devillez, A., Lecolier, G., and Billaudel, P. (2002). Recursive learning in real time using fuzzy pattern matching. *Mathematics and Computers in Simulation*, 60:209–216.

Su, M.-C. and Liu, Y.-C. (2005). A new approach to clustering data with arbitrary shapes. *Pattern Recognition*, 38:1887–1901.

Traore, M., Duviella, E., and Lecoeuche, S. (2009). Dynamical clustering technique to estimate the probability of the failure occurrence of process subjected to slow degradation. In *ICINCO, Milan, Italy*, pages 636–643.

Weihua, L., Harigopal, R., and Sirish, S. (2003). Subspace identification of continuous time models for process fault detection and isolation. *Journal of Process Control*, 13:407–421.

Xie, L., Soh, Y. C., and de Souzi, C. E. (1994). Robust kalman filtering for uncertain discrete-time systems. *IEEE Transaction on Automatic Control*, 93:131 01314.