Sensor fault diagnosis of inland navigation system using physical model and pattern recognition approach

K. Horváth * **J. Blesa** *** **E. Duviella** * **L. Rajaoarisoa** * **V. Puig** ***
K. Chuquet **

* *Mines-Telecom Institute, Mines Douai, France (e-mail:*
{eric.duviella,klaudia.horvath,lala.rajaoarisoa}@mines-douai.fr).
** *VNF - Service de la navigation du Nord Pas-de-Calais, France (e-mail:*
karine.chuquet@vnf.fr)
*** *Institut de Robòtica i Informàtica Industrial, CSIC-UPC. Barcelona,*
Spain (e-mail:{joaquim.blesa,vicenc.puig}@upc.edu)

Abstract: In recent years, inland navigation networks benefit from the innovation of the instrumentation and SCADA systems. These data acquisition and control systems lead to a reactive asset-management of inland navigation networks. However, sensors and actuators are subject to faults due to the strong effects of the environment, aging, etc. In this paper, a sensor Fault Detection and Isolation (FDI) approach is proposed using an Integrator-Delay-Zero (IDZ) model, interval observers and the dynamic classification algorithm *AUDyC*. The combined use of these approaches allows the improvement of the sensor fault diagnosis. The proposed approach is introduced through the case study of the Cuinchy-Fontinettes reach in the north of France.

Keywords: Fault diagnosis, Model-based, Pattern recognition, Large scale system, Inland navigation.

1. INTRODUCTION

To improve the management of the inland navigation networks, national companies are still modernizing the data acquisition and remote control systems. They benefit from the progress in new technologies for sensors and actuators (Fastenbauer et al., 2007) and SCADA (Supervisory Control and Data Acquisition) systems (Malaterre and Chateau, 2007). The data that is taken into account to supervise the networks and to control the actuators have to be safe in order to avoid serious failures of the navigation and the infrastructure. Thus, Fault Detection and Isolation (FDI) methods have to be designed and implemented. The FDI techniques are generally divided into two families: model-based and data-driven approaches.

Model-based approaches require accurate models of the system dynamics. A survey of model-based techniques is proposed in Frank et al. (2000). Model-based Fault diagnosis techniques have recently addressed the supervision of open flow channels. Parity equations method in Blesa et al. (2010), the cumulative sum (CUSUM) algorithm applied to model-based residuals in Bedjaoui and Weyer (2011), physical and nonlinear black-box models of a real undershot/overshot gate in Pocher et al. (2012), the evaluation of model-based residuals Nabais et al. (2013) and subspace identification in Akhenak et al. (2013).

The second family gathers the data-driven approaches that have been recently surveyed in Qin (2012). These methods perform the fault diagnosis by directly processing the signals collected from the sensors (Hartert et al., 2010). Amongst these methods, pattern recognition techniques consist in representing the operating modes of the system, *i.e.* normal and faulty, by classes. Some dynamical classification algorithms are proposed in the literature such as the CDL algorithm (Cluster Detection and Labeling), algorithms based on adaptive resonance theory (ART) networks (Eltoft and de Figueiredo, 1998), or AUDyC (Auto-Adaptive Dynamical Clustering) algorithm (Lecoeuche et al., 2004). The characteristics of the classes are followed online and faults can be detected using metrics between classes. A FDI technique based on the algorithm *AUDyC* is proposed in Pocher et al. (2012) to detect sensor faults of a navigation reach only from real data.

A FDI strategy for sensors of inland navigation reaches, that combines model-based and data-driven approaches, is proposed in Duviella et al. (2013b). The model is a gray-box model dedicated to multi-input/multi-output (MIMO) and time-delay open-flow channel systems (Duviella et al., 2013a). It is applied to the navigation reach of the north of France. However in this paper, the residuals from gray-box model are ideal because they are affected only by faults. In real applications, the presence of disturbances, noise and modelling errors causes the residuals to become nonzero even in the absence of faults. Therefore, the fault detection procedure must be robust against these undesired effects (Chen and Patton, 1999). When parametric uncertainties are taken into account, the healthy system model should include a vector of uncertain parameters bounded by a set that contains all possible parameter values of normal operation. In the robust fault detection literature, so far, parameters have been bounded using intervals and the resulting model is known as an interval model (Puig et al., 2008). When a dynamic system is modelled using an interval model, the predicted output is also usually bounded by an interval. Then, fault detection test is based on checking if zero is contained or not in the residual interval after propagating the parameter uncertainty. In this paper, the FDI approach proposed in Duviella et al. (2013b) is improved by using an interval method based on a white-box model of
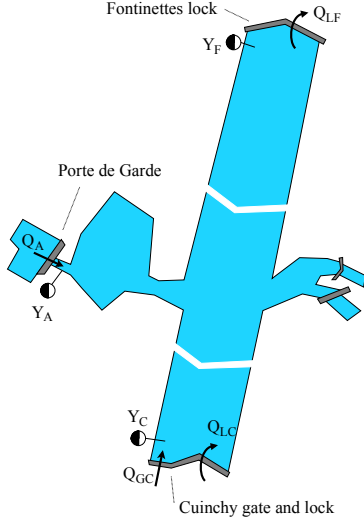
Fig. 1. Schematics of the CFR

navigation reach: the Integrator-Delay-Zero (IDZ) model. The nominal model and the interval outputs provided by the interval model are used as features for the classification algorithm *AUDyC* that permits the detection and isolation of sensor faults. The proposed approach is applied to a navigation reach in the north of France.

The structure of the paper is the following: Section 2 is dedicated to the description of the IDZ model. The interval model is proposed in Section 3. In Section 4, the algorithm *AUDyC* and the fault detection module are presented. The proposed approach is applied to the real network in Section 5.

2. CUINCHY-FONTINTETTES CASE STUDY

2.1 Description

The Cuinchy-Fontintettes reach (CFR) belongs to the navigation network of the north of France. The system has to ensure the waterway between Paris and the major ports of the North Sea. The CFR reach is $42\ km$ long being bounded by the lock of Cuinchy at the upstream and by the lock of Fontinettes at the downstream end (*see* Figure 1). There are two additional inflows to the reach: (1) at the upstream end on the side of the lock of Cuinchy there is a gate (Q_{GC}), (2) in the middle of the reach the river Aire is connected to the CFR by a gate (Q_A). The water level is mainly influenced by the operations of the two locks: when lock Cuinchy is operated it releases its water volume ($3600\ m^3$) to the CFR, when lock Fontinettes is operated it takes water ($25000\ m^3$) from the CFR. As the volume of the lock of Fontinettes is considerably bigger than the volume of Cuinchy, it produces a negative wave of amplitude about 15 cm that is able to travel up and down in the reach.

In order to model the reach, the discharge of the lock of Cuinchy (Q_{LC}) and the gate at the side (Q_{GC}) is considered together, and denoted as Q_C. The other two inputs to the model are the discharge in the middle of the canal, at Aire (Q_A) and the discharge at the downstream end from the operation of lock Fontinettes (Q_F). The three measured variables are the water levels at the three locations of the canal: at the upstream end, at Cuinchy (Y_C), in the middle of the reach at Aire (Y_A) and at the downstream end of the reach at lock Fontinettes (Y_F).

2.2 The IDZ model

The dynamics of the open-flow channel can be described by the Saint-Venant (SV) equations that are hyperbolic partial differential equations, (Chow, 1959). These equations are highly non-linear and they have no analytical solution in the general case. It is common to use their numerical solution for simulation purposes, however for other goals (as real-time control or fault detection) these solutions are too complex.

For this reason there are several simplified models that have been deduced from the SV equations with different simplifications, for example Hayami (Hayami, 1951), ID (Schuurmans, 1995), IR (van Overloop et al., 2010) and IDZ (Litrico and Fromion, 2004). The IDZ model is based on the linearisation of the SV equations around a set-point, using three points of a spatial discretization the canal reach is divided into two parts: an upstream (uniform flow) part and a downstream (backwater) part. Then, the approximation is based on three points: the upstream and the downstream ends and one intermediate point.

The model contains four transfer functions

$$\begin{bmatrix} Y_{up} \\ Y_{down} \end{bmatrix} = \begin{bmatrix} p_{11}(s)\ p_{12}(s) \\ p_{21}(s)\ p_{22}(s) \end{bmatrix} \begin{bmatrix} Q_{up} \\ Q_{down} \end{bmatrix} \quad (1)$$

where $p_{11}(s)$ relates the upstream discharge (Q_{up}) and the downstream water level (Y_{down}), $p_{12}(s)$ relates the downstream discharge (Q_{down}) and the upstream water level (Y_{up}), $p_{21}(s)$ relates the upstream discharge and the upstream water level and $p_{22}(s)$ relates the downstream discharge and the downstream water level. The structure of each transfer function is the same. They contain an integrator, a delay and a zero as the name of the model indicates (*see* Eq. (2)).

The integrator is responsible for the modelling at low frequencies, where the canal pool acts as an integrator, that is the water level increases with the integral value of the discharge, like a tank, multiplied by the inverse of the surface of the canal reach, known as the backwater area, A_s. This area is close to the product of the width and the length of the canal.

The delay is the time it takes for a wave to travel from the point of disturbance to the point of measurement. It is slightly different in the upstream (τ_u) and in the downstream (τ_d) directions. Their approximation is detailed in Litrico and Fromion (2004).

The IDZ model contains a zero, that accounts for the first fast increase in water level after a disturbance occurs. The general form of the transfer functions can be written as:

$$G(s) = \frac{p_1 s + 1}{A_s s} e^{-\tau s}, \quad (2)$$

where the parameter p_1 accounts for the zero and τ can be τ_u or τ_d depending of the direction of the disturbance.

2.3 Application to the Cuinchy-Fontinettes case study

The IDZ modelling approach is applied to the Cuinchy-Fontinettes case study considering that it can be decomposed in three reaches that are described by the parameters presented in Table 1. The Manning's coefficient was chosen to be 0.035 based on Chow (1959). As the bed slope is very small (almost zero), the reach is affected by backwater and therefore the whole reach was modelled just one part as backwater part. The CFR system has three inputs (Q_C, Q_A, Q_F) and three outputs

Name	Length (m)	Width (m)	Depth (m)	Discharge (m^3/s)	Bottom slope (-)
CF	42300	52	3.8	3	2.5×10^{-7}
CA	28700	52	3.8	0	2.5×10^{-7}
AF	13600	52	3.8	0	2.5×10^{-7}

Table 1. Physical data of the CFR

Name	Backwater surface (m^2)	τ (s)	p_1 (-)	TF	Reach
G_{CC}	2200 $\times 10^3$	0	13990	p_{11}	CF
G_{AC}	2200 $\times 10^3$	4711	4885	p_{12}	CA
G_{FC}	-2200 $\times 10^3$	6943	3969	p_{12}	CF
G_{CA}	2197 $\times 10^3$	4688	4892	p_{21}	CA
G_{AA}	2197 $\times 10^3$	0	10380	p_{22}	CA
G_{FA}	-2200 $\times 10^3$	2233	5975	p_{12}	AF
G_{CF}	2197 $\times 10^3$	6909	3969	p_{21}	CF
G_{AF}	2197 $\times 10^3$	2222	5979	p_{21}	AF
G_{FF}	-2197 $\times 10^3$	0	9196	p_{22}	CF

Table 2. Properties of the three reaches

$(Y_C,\ Y_A,\ Y_F)$. Each output can be modelled by using three transfer functions with the structure of (2):

$$\begin{cases} Y_C(s) = G_{CC}(s)Q_C(s) + G_{AC}(s)Q_A(s) + G_{FC}(s)Q_F(s) \\ Y_A(s) = G_{CA}(s)Q_C(s) + G_{AA}(s)Q_A(s) + G_{FA}(s)Q_F(s) \\ Y_F(s) = G_{CF}(s)Q_C(s) + G_{AF}(s)Q_A(s) + G_{FF}(s)Q_F(s) \end{cases} \quad (3)$$

The transfer functions show the effect of the discharges to the water levels, *e.g.* G_{CF} is the transfer function between the discharge at Fontintettes and the level at Cuinchy.

The modelling of the intermediate discharge in the middle of the reach was not that straightforward. In order to model the transfer functions between Cuinchy and Aire and Aire and Fontinettes, all the IDZ parameters (the time delay, the zero and the backwater surface) were calculated following Litrico and Fromion (2004). Finally, the backwater surface was substituted with the one calculated for the whole canal: the discharge that enters in the middle of the reach contributes to the water level of the whole reach. The calculated parameters are summarized in Table 2. The following remarks should be considered:

(1) Note that the all the backwater areas are very similar, but there is a slight difference between them in the upstream and downstream directions. However, they are similar between the reaches.

(2) The sign of the backwater area is negative belonging to the effect of the discharge at Fontinettes, that is the downstream end. The reason is that this discharge leaves the canal, therefore its effect is decreasing the water level.

(3) The obtained time delays are similar, but not equal in the upstream and downstream directions: *e.g.* in the downstream direction from Cuinchy to Aire is $4688s$ (line 4), while in the upstream direction form Aire to Cuinchy is $4711s$ (line 2). The reason is that the wave travelling with the direction of the celerity in the downstream direction the velocity of the bulk of the water is added to the wave celerity, while when travelling upstream this velocity slows the wave down. However, as it can be seen that the two values are very close to each other, the velocity of the bulk of the water is negligible compared to the wave celerity.

3. RESIDUAL GENERATION USING INTERVAL OBSERVERS

3.1 Interval observers

As discussed in the introduction, the FDI approach proposed in this paper for CFR system aims at combining model-based and data-based approaches.

The model-based approach will be used to generate residuals and their thresholds by means of interval observer methodology (Puig et al., 2008). With this objective, the CFR model (3) is expressed in discrete-time state space form as follows:

$$x(k+1) = A(\tilde{\theta})x(k) + B(\tilde{\theta})u(k) \quad (4)$$
$$y(k) = C(\tilde{\theta})x(k) + F_y(\tilde{\theta})f_y(k) + \tilde{v}(k) \quad (5)$$

where $u(k) \in \mathbb{R}^{n_u}$ is the system input, $y(k) \in \mathbb{R}^{n_y}$ is the system output, $x(k) \in \mathbb{R}^{n_x}$ is the state-space vector, $\tilde{v}(k) \in \mathbb{R}^{n_y}$ is the output noise that is assumed to be bounded $|\tilde{v}_i(k)| < \sigma_i$ with $i = 1, \ldots, n_y$, $f_y(k) \in \mathbb{R}^{n_y}$ represents faults in the output sensors. $A(\tilde{\theta})$, $B(\tilde{\theta})$, $C(\tilde{\theta})$ and $F_y(\tilde{\theta})$ are matrices of appropriate dimensions where $\tilde{\theta} \in \mathbb{R}^{n_\theta}$ is the vector of real parameters.

Then, the system (4)-(5) can be monitored using a linear observer with Luenberger structure that uses an *interval model* of the system, *i.e.* a model with parameters θ bounded by intervals:

$$\theta \in \Theta = \left\{ \theta \in \mathbb{R}^{n_\theta} \mid \underline{\theta}_i \leq \theta_i \leq \bar{\theta}_i, i = 1, \ldots, n_\theta \right\} \quad (6)$$

that represent the uncertainty about the exact knowledge of the real parameters $\tilde{\theta}$. This observer, known as an *interval observer*, is expressed as follows (Meseguer et al., 2010):

$$\hat{x}(k+1, \theta) = (A(\theta) - LC(\theta))\hat{x}(k, \theta) + B(\theta)u(k) + Ly(k)$$
$$= A_0(\theta)\hat{x}(k, \theta) + B(\theta)u(k) + Ly(k) \quad (7)$$
$$\hat{y}(k, \theta) = C(\theta)\hat{x}(k, \theta)$$

where $\hat{x}(k, \theta)$ is the estimated system state vector, $\hat{y}(k, \theta)$ is the estimated system output vector and $A_0(\theta) = A(\theta) - LC(\theta)$ is the observer matrix.

The observer gain matrix $L \in \mathbb{R}^{n_x \times n_y}$ is designed to stabilize the matrix $A_0(\theta)$ and to guarantee a desired performance regarding fault detection for all $\theta \in \Theta$ using the LMI pole placement approach (Chilali and Gahinet, 1996).

The input/output form of the system (4)-(5) using the shift operator q^{-1} and assuming zero initial conditions is given by:

$$y(k) = y_0(k, \tilde{\theta}) + G_{f_y}(\tilde{\theta})f_y(k) + \tilde{v}(k) \quad (8)$$

where $y_0(k, \tilde{\theta})$ is the system output when the system is not affected by faults, disturbances and noises:

$$y_0(k, \tilde{\theta}) = G_u(q^{-1}, \tilde{\theta})u(k) \quad (9)$$
$$G_u(q^{-1}, \tilde{\theta}) = C(\tilde{\theta})(qI - A(\tilde{\theta}))^{-1}B(\tilde{\theta}) \quad (10)$$
$$G_{f_y}(\tilde{\theta}) = F_y(\tilde{\theta}) \quad (11)$$

The input/output form of the observer (7) is expressed as follows:

$$\hat{y}(k, \theta) = G(q^{-1}, \theta)u(k) + H(q^{-1}, \theta)y(k) \quad (12)$$

with:

$$G(q^{-1}, \theta) = C(\theta)(qI - A_0(\theta))^{-1}B(\theta) \qquad (13)$$

$$H(q^{-1}, \theta) = C(\theta)(qI - A_0(\theta))^{-1}L \qquad (14)$$

The effect of the uncertain parameters θ on the observer temporal response $\hat{y}(k, \theta)$ will be bounded using an interval satisfying:

$$\hat{y}(k, \theta) \in \left[\underline{\hat{y}}(k), \overline{\hat{y}}(k)\right] \qquad (15)$$

Such interval can be computed independently for each output $i = 1, \ldots, n_y$, neglecting couplings among outputs, as follows:

$$\underline{\hat{y}_i}(k) = \min_{\theta \in \Theta} \hat{y}_i(k, \theta) \quad \text{and} \quad \overline{\hat{y}}_i(k) = \max_{\theta \in \Theta} \hat{y}_i(k, \theta) \qquad (16)$$

subject to the observer equations given by (7). The optimization problems (16) could be solved using numerical methods as in Puig et al. (2003).

Finally, taking into account that the additive noise in the system (5) is bounded, the following condition should be satisfied in a non-faulty scenario:

$$y_i(k) \in \left[\underline{\hat{y}_i}(k) - \sigma_i, \overline{\hat{y}_i}(k) + \sigma_i\right] i = 1, \ldots, n_y \qquad (17)$$

3.2 Parameter uncertainty estimation

One of the key points in passive robust model-based fault detection is how models and their uncertainty bounds are obtained. Recently, some methodologies that provide a model with its uncertainty have been developed, but always thinking of its application to control (Reinelt et al., 2002). One of the methodologies assumes the bounded and unknown description of the noise and parametric uncertainty. This methodology is known as *bounded-error* or *set-membership estimation* (Milanese et al., 1996), which produces a set of parameters consistent with the selected model structure and the pre-specified noise bounds. This approach is used for estimating parametric uncertainty of the interval observers in (7).

Regarding the uncertain variables that appear in (7), it is assumed that a priori theoretical or practical considerations allow to obtain useful intervals associated to measurement noises, leading to an estimation of the noise bound σ. The goal of the parameter estimation algorithm is to characterize the parameter set Θ (here a box) consistent with the data collected in a fault-free scenario. Given N measurements of system inputs $u(k)$ and outputs $y(k)$ from a scenario free of faults and rich enough from the identifiability point of view, and a nominal model described by a vector θ_n that can be obtained by the physical knowledge of the system or using a standard least-square parameter estimation algorithm Ljung (1987), the uncertain parameter estimation algorithm proceeds by solving the following optimization problem that can be solved by bi-level optimization (Bard, 1988):

$$\begin{aligned}
&\min \ \alpha \\
&\text{subject to :} \\
&\quad y_i(k) \in \left[\underline{\hat{y}_i}(k) - \sigma_i, \overline{\hat{y}_i}(k) + \sigma_i\right] \quad i = 1, \ldots, n_y \quad k = 1, \ldots, N \\
&\quad \underline{\hat{y}_i}(k) = \min_{\theta \in \Theta} \hat{y}_i(k, \theta) \quad i = 1, \ldots, n_y \quad k = 1, \ldots, N \\
&\quad \overline{\hat{y}}_i(k) = \max_{\theta \in \Theta} \hat{y}_i(k, \theta) \quad i = 1, \ldots, n_y \quad k = 1, \ldots, N \\
&\quad \hat{y}(k, \theta) = G(q^{-1}, \theta)u(k) + H(q^{-1}, \theta)y(k) \quad k = 1, \ldots, N \\
&\quad \Theta = [\theta_n(1 - \alpha), \theta_n(1 + \alpha)]
\end{aligned}$$
$$(18)$$

3.3 Residual generation

Fault detection is based on generating a nominal residual comparing the measurements of physical system variables $y(k)$ with

their estimation $\hat{y}(k)$ provided by the observer (7):

$$r^o(k) = y(k) - \hat{y}(k, \theta_n) \qquad (19)$$

where $r^0(k) \in \mathbb{R}^{n_y}$ is the nominal residual. According to Gertler (1998), the computational form of the nominal residual generator, obtained using (7), is:

$$r^o(k) = \left(I - H(q^{-1}, \theta_n)\right)y(k) - G(q^{-1}, \theta_n)u(k) \qquad (20)$$

that has been derived taking into account the input/output form of the observer (12).

When considering model uncertainty located in parameters, the residual generated by (19) will not be zero, even in a non-faulty scenario. To cope with the parameter uncertainty effect, a passive robust approach based on adaptive thresholding can be used (Puig et al., 2006). Thus, using this passive approach, the effect of parameter uncertainty in the components $r_i(k)$ of residual $r(k)$ (associated to each system output $y_i(k)$) is bounded by the interval (Puig et al., 2003):

$$r_i^o(k) \in [\underline{r}_i(k) - \sigma_i, \overline{r}_i(k) + \sigma_i] \quad i = 1, \ldots, n_y \qquad (21)$$

where:

$$\underline{r}_i(k) = \underline{\hat{y}_i}(k) - \hat{y}_i(k, \theta_n) \text{ and } \overline{r}_i(k) = \overline{\hat{y}_i}(k) - \hat{y}_i(k, \theta_n) \qquad (22)$$

where $\underline{\hat{y}}_i(k)$ and $\overline{\hat{y}}_i(k)$ are the bounds of the system output estimation computed component-wise using the interval observer (7) and obtained according to (16).

4. FAULT DIAGNOSIS USING DYNAMICAL CLASSIFIER *AUDYC* AND RESIDUALS

4.1 Overview

The fault diagnosis approach proposed in this paper is based on using residuals and their uncertainty intervals presented in previous section combined with a dynamical classifier. The idea is that this classifier is in charge of tracking the evolution of the system residuals in the residual space. This allows a more sensitive, faster detection than using just thresholds. An additional motivation is the specific characteristic of this water system: the measurement can change fast within short time. Therefore with only using the residuals and threshold it would be either not sensitive enough or would lead to false alarm. The use of classifiers can compensate for this fast movement, and acts as a filter by finding the change of the class.

When the residuals leaves the region of normal behaviour established by the uncertainty bounds a fault is detected. Fault isolation is achieved when residuals enter in a region of the residual space corresponding to a given fault signature. The regions are created by training the classifier using faulty and non-faulty scenarios created in simulation.

4.2 AUDyC classifier

The classifier *AUDyC* aims at characterizing on line the free-fault and faulty modes of the system by Gaussian classes according to pertinent indicators, *i.e.* the features. A *Representation Space* is built according to the features, where the classes are represented. The class of the free-fault mode, denoted C_n, is determined with fault-free data. Then, the characteristics of the class, *i.e.* its center and its covariance matrix, are updated with each new data. This evolutionary class is denoted C_k.

The update of the classes parameters is performed recursively on a sliding window of size N_{fen} for each new observation

$X_k = (x_1, x_2, ... x_n)$, with n the number of features. The principle of the procedure is detailed in Lecoeuche et al. (2004). To detect a modification of the characteristics of the evolutionary class, the Kullback-Leiber distance between C_n and C_k is computed at each sample time. It is denoted $d_{Kl}(M_k, M_n)$ and expressed as:

$$\begin{cases} d_{Kl}(M_k, M_n) = & 1/2(M_k - M_n)^T(\Sigma_k^{-1} + \Sigma_n^{-1})(M_k - M_n) + \\ & 1/2 trace(\Sigma_k^{-1}\Sigma_n^{-1} + \Sigma_k + \Sigma_n^{-1}) - D, \end{cases}$$

(23)

where M_k is the center of the class C_k and M_n the center of the normal class C_n, Σ_k and Σ_n the covariance matrices of the classes C_k and C_n respectively, D the dimension of the *Representation Space*, and $trace$ computes the sum of diagonal elements of the matrix.

When a fault occurs, the characteristics of the evolutionary class C_k changes. The class C_k moves in the *Representation Space* from the normal class C_n and its shape (i.e. Σ_k) can be modified. These modifications have an impact on the Kullback-Leiber distance $d_{Kl}(M_k, M_n)$. Then, a sensor fault is detected when the $d_{Kl}(M_k, M_n)$ crosses a prefixed threshold th. This threshold is tuned according to data or expert knowledge on the system.

4.3 Application to the Cuinchy-Fontinettes case study

Considering the CFR system, there are three sensors that can be subject to faults. Using the IDZ models associated to the interval models, three features are defined for each sensor by means of the residuals and their intervals generated using the corresponding interval observers presented in previous section: *i.e.* $r_C(k)$, $\underline{r}_C(k)$ and $\overline{r}_C(k)$ for the sensor in Cuinchy Y_C, $r_A(k)$, $\underline{r}_A(k)$ and $\overline{r}_A(k)$ for the sensor in Aire Y_A, finally, $r_F(k)$, $\underline{r}_F(k)$ and $\overline{r}_F(k)$ for the sensor in Fontinettes Y_F. The residuals $r_i(k)$ are directly issued from the nominal IDZ model and the measurements. Thus, three *Representation Spaces* are defined for the CFR.

5. RESULTS

A numerical model of the CFR was built using the SIC (Simulation of Irrigation Canals) (Malaterre, 2006) software. SIC uses a finite difference method to solve the SV equations implicitly. The cross sections were placed at about each $100\ m$, using surveyed data. The discretization time was chosen in order the Courant-Friedrichs-Levy condition to be kept. After this validation, the numerical simulator is used to reproduce the scenarios presented in the following.

The transfer functions were discretized by using zero order hold, with a sampling time of $300s$. The models of the three water levels were compared to the results of the distributed solution of the SV equations (by the SIC software) and the bed slope of the IDZ model was adjusted to represent better the real data.

According to the IDZ model described in Section 2, the real behaviour of the CFR system in presence of sensor faults and additive error noise can be expressed in state space as (4)-(5) with $y = (Y_C(k), Y_A(k), Y_F(k))$ and $u = (Q_C(k), Q_A(k), Q_F(k))$.

Nominal parameters θ_n where obtained as explained in Section 2.3 using a fault free scenario rich enough under the identifi-

cation point of view. The interval observer is calibrated as it was described in Section 3. The scenario used for calibration corresponds to normal operation of the canal reach: about 10 lock operations per day both upstream and downstream and changing flow input in the middle of the reach.

As can be deduced from the IDZ equations (3), every sensor level fault affects only the residuals associated to its level. Then, the isolation is trivial: when an inconsistency is detected in one residual, a fault in the level associated to the residual is concluded.

Once the interval model has been calibrated, different faults have been simulated in order to verify the effectiveness in fault detection using directly (21) and using the AUDYC classifier with the nominal, maximum and minimum residuals.

Figure 2 shows the measured level in Cuinchy and interval bounds in a fault scenario of -2cm offset level applied at time $t = 5000$ minutes. It shows the evolution of the Cuinchy level (calculated by the SIC simulator) for six days long. The interval bounds computed by (16). It can be seen when the fault occurs the "measurement" leaves the interval.
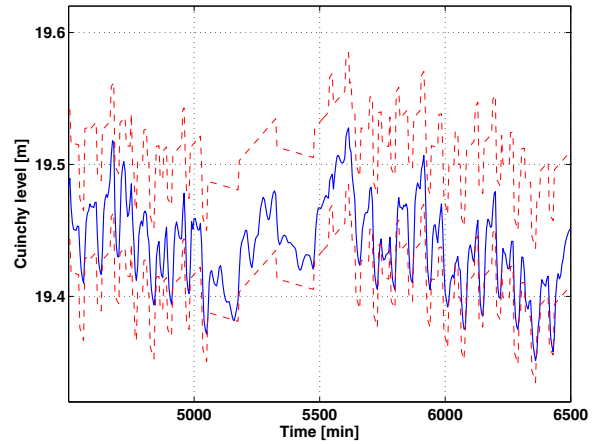
Fig. 2. Cuinchy level (blue line) and bounds (red dashed line) and Fault detection test in Fault scenario.

The Kullback-Leiber distance $d_{Kl}(M_k, M_n)$ between these two classes are determined on line and shown in Figure 3, in continuous blue line for a free-fault scenario and in dashed red line for fault scenario. The threshold is tuned as 0.08. With this threshold, a fault of $2cm$ (0.5%) is detected within 23 samples (2 hours). Afther the detection there are two intervals of no detection, while the fault is present.

6. CONCLUSION AND PERSPECTIVES

A sensor fault diagnosis method is proposed based on the combination of model-based approach and dynamic classification. A white box model of the system is used in combination with an interval observer to generate residuals for the classifier. The method is successfully applied to a case study of an inland navigation reach. This result shows that it is a promising line of research to combine the model-based and the data-driven FDI methods.
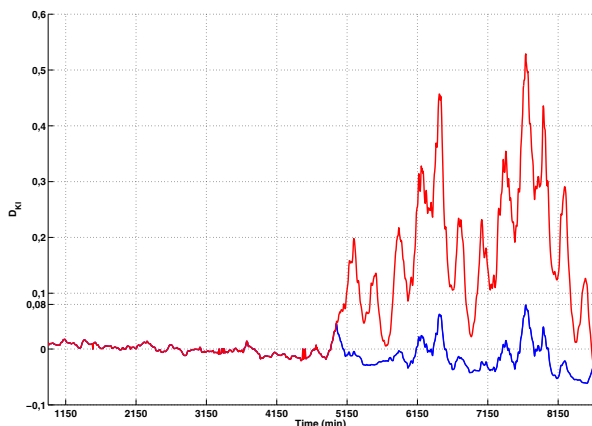
Fig. 3. Kullback-Leiber distance $d_{Kl}(M_k, M_n)$ between these two classes - free-fault in blue, fault in red.

ACKNOWLEDGMENT

REFERENCES

A. Akhenak, E. Duviella, L. Bako, and S. Lecoeuche. Online fault diagnosis using recursive subspace identification algorithm of a dam-gallery open channel system. *Control Engineering Practice*, In press, 2013.

J.F. Bard. *Practical Bilevel Optimization: Algorithms and Applications*. The Netherlands: Kluwer, 1988.

N. Bedjaoui and E. Weyer. Algorithms for leak detection, estimation, isolation and localization in open water channels. *Control Engineering Practice*, 19(6):564–573, 2011.

J. Blesa, V. Puig, and Y. Bolea. Fault detection using interval LPV models in an open-flow canal. *Control Engineering Practice*, 18(5):460 – 470, 2010.

J. Chen and R. J. Patton. *Robust Model-based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers, 1999.

M. Chilali and P. Gahinet. H_∞ Design with pole placement constraints: an LMI approach. *IEEE Transactions on Automatic Control*, 41(3):358–367, 1996.

V. T. Chow. *Open-channel hydraulics*. McGraw-Hill Book Co. Inc, New York, 1959.

E. Duviella, J. Blesa, L. Bako, Y. Bolea, M. Sayed-Mouchaweh, V. Puig, and K. Chuquet. Inland navigation channel model: Application to the cuinchy-fontinettes reach. In *The 10th IEEE International Conference on Networking, Sensing and Control*, Paris, France, April 10-12. 2013a.

E. Duviella, L. Rajaoarisoa, J. Blesa, and K. Chuquet. Fault detection and isolation of inland navigation channel: Application to the Cuinchy-Fontinettes reach. In *CDC 52nd IEEE Conference on Decision and Control*, December 2013b.

T. Eltoft and R.I.P. de Figueiredo. A new neural network for cluster-detection-and-labeling. *Neural Networks, IEEE Transactions on*, 9(5):1021–1035, 1998.

M. Fastenbauer, M. Sattler, and G. Schilk. River information services for commercial users in the inland waterway sector. In *LINDI 2007. International Symposium on Logistics and Industrial Informatics*, pages 31–36, 2007.

P. M. Frank, S. X. Ding, and B. Köppen-Seliger. Current developments in the theory of FDI. In A. Edelmayer, editor, *SAFEPROCESS*, pages 16–17, 2000.

J. Gertler. *Fault Detection and Diagnosis in Engineering Systems*. Taylor & Francis, 1998.

L. Hartert, M. Sayed Mouchaweh, and P. Billaudel. Monitoring of non stationary systems using dynamic pattern recognition. In *Intelligent Industrial Systems: Modelling, Automation and Adaptive Behaviour*. IGI, 2010.

S. Hayami. On the propagation of flood waves. *Bulletins-Disaster Prevention Research Institute, Kyoto University*, 1: 1–16, 1951.

S. Lecoeuche, C. Lurette, and S. Lalot. New supervision architecture based on on-line modelling of non-stationary data. *Neural Comput. Appl.*, 13(4):323–338, December 2004.

X. Litrico and V. Fromion. Simplified modeling of irrigation canals for controller design. *Journal of Irrigation and Drainage Engineering*, 130(5):373–383, 2004.

L. Ljung. *System Identification: Theory for the User*. Prentice Hall Information and System Sciences Series, 1987.

P.-O. Malaterre. Sic 5.20, simulation of irrigation canals, 2006. URL http://www.cemagref.net/sic/sicgb.htm.

P.-O. Malaterre and C. Chateau. Scada interface of the sic software for easy real time application of advanced regulation algorithms. In *Second Conference on SCADA and Related Technologies for Irrigation System Modernization*, 2007.

J. Meseguer, V. Puig, T. Escobet, and J. Saludes. Observer gain effect in linear interval observer-based fault detection. *Journal of Process Control*, 20:944–956, 2010.

M. Milanese, J. Norton, H. Piet-Lahanier, and É. Walter, editors. *Bounding Approaches to System Identification*. Springer, 1996.

J. Lemos Nabais, L. F. Mendona, and M. Ayala Botto. A multi-agent architecture for diagnosing simultaneous faults along water canals. *Control Engineering Practice*, (0):–, 2013.

O. Le Pocher, E. Duviella, L. Bako, and K. Chuquet. Sensor fault detection of a real undershot/overshot gate based on physical and nonlinear black-box models. In *SAFEPROCESS*, 2012.

V. Puig, J. Saludes, and J. Quevedo. Worst-Case Simulation of Discrete Linear Time-Invariant Interval Dynamic Systems. *Reliable Computing*, 9(4):251–290, 2003.

V. Puig, A. Stancu, T. Escobet, F. Nejjari, J. Quevedo, and R. J. Patton. Passive Robust Fault Detection using Interval Observers: Application to the DAMADICS Benchmark Problem. *Control Engineering Practice*, 14(6):621–633, 2006.

V. Puig, J. Quevedo, T. Escobet, F. Nejjari, and S. de las Heras. Passive Robust Fault Detection of Dynamic Processes using Interval Models. *IEEE Transactions on Control Systems Technology*, 16:1083–1089, 2008.

S. J. Qin. Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 36(2):220 – 234, 2012.

W. Reinelt, A. Garulli, and L. Ljung. Comparing different approaches to model error modeling in robust identification. *Automatica*, 38(5):787 – 803, 2002.

J. Schuurmans. Open-channel flow model approximation for controller design. *Applied Mathematical Modelling*, 19(9): 525–530, 1995.

P.-J. van Overloop, I. J. Miltenburg, X. Bombois, A. J. Clemmens, R. Strand, and N. van de Giesen. Identification of resonance waves in open water channels. *Control Engineering Practice*, 18(8):863–872, 2010.