

Temporal/Spatial Model-Based Fault Diagnosis vs. Hidden Markov Models Change Detection Method: Application to the Barcelona Water Network

J. Quevedo, C. Alippi, M.À. Cugueró, S. Ntalampiras, V. Puig, M. Roveri, D. García

Abstract—This paper deals with a comparison of two different fault diagnosis frameworks. The first method is based on a temporal/spatial model-based analysis by exploiting *a-priori* information about the system under study, so fault detection is based on monitoring the residuals of combined spatial and time series models obtained from the network. The second method aims at characterizing and detecting changes in the probabilistic pattern sequence of data coming from the network. Relationships between data streams are modelled through sequences of linear dynamic time-invariant models whose trained coefficients are used to feed a Hidden Markov Model (HMM). When the pattern structure of incoming data cannot be explained by the trained HMM, a change is detected. Here, the performance obtained from this two distinct approaches is examined by using a dataset coming from the Barcelona water transport network.

Index Terms—fault diagnosis, sensor networks, pattern recognition, time series models.

I. INTRODUCTION

Critical Infrastructure Systems (CIS) including water, gas or electrical networks are complex large-scale systems which need of highly sophisticated supervisory and control schemes to satisfy a certain degree of performance when unfavorable conditions are occurring. To deal with such a problem, the use of a fault detection and isolation (FDI) system which is capable to detect and correct these faults by activating fault tolerance control (FTC) mechanisms is highly desirable. Hence, the FDI method aims to identify which is the actual fault occurring in the system. According to the literature, the problem of FDI may be generally addressed by two main strategies: hardware and analytical redundancy. The first approach is based on the use of extra sensors, whilst the second one is based on the use of software sensors or models combined with information gathered by the sensor measurements. Even though hardware redundancy is desirable in critical systems, the use of the latter in large-scale systems may be pretty expensive, because of the calibration and maintenance actions to be performed on the system.

In this work, two different fault diagnosis approaches will be studied:

- The first approach is based on checking the consistency between the observed and the nominal system behavior, by means of a set of analytical redundancy relations (ARRs) which relate the measured system variables with a model of normal (faultless) operation of the monitored system. When an inconsistency is detected, the fault

detection and isolation mechanism is activated in order to detect and identify the possible fault (?).

- The motivation behind the second approach is the fact that the relationships among different sensor measurements monitoring the same environmental phenomenon follow a statistical pattern over time. Differently, a change in the process or a fault in sensor gathering data might occur and this would result in a change of statistical pattern. This method aims at modelling the normal state by means of a Hidden Markov Model (HMM) trained on the parameters of linear dynamic time-invariant models (Alippi et al., 2012). A finite state machine representing the fault-free state is obtained which is then applied onto incoming data. A log-likelihood is derived which may be considered as a measure of the statistical similarity between the incoming data and the data trained by the HMM in an initial fault-free situation. In case the likelihood decreases below a predefined threshold a change is detected.

To evaluate the FDI methodologies in CIS, the Barcelona drinking water network is used as the case study in this paper. In ideal situations, the knowledge of the analytical model as considered in the first approach should lead to the optimal solution. However, it may be noted that analytical models may be affected by several practical aspects of the system, such as the potential uncertainty on the model parameters (e.g. usable tanks surface), the difficulty to have an on-line well-calibrated model due to frequent network topology changes (caused by e.g. new elements like tanks added or blocked pipes as a result of maintenance operations) and common changes on the consumers demand behavior which is hard to determine in real-time operation. Hence, a probabilistic approach as suggested in the second method is also a useful and effective alternative to the use of analytical models.

Previous works has been realized performing fault diagnosis in this water network, e.g. in (?), where validation (detection) and reconstruction of the missing and false data of the flowmeters installed in the network is applied, in (Quevedo et al., 2011), where a leakage localization method based on the pressure measurements and pressure sensitivity analysis of network nodes is presented, or in (Nejjari et al., 2012), where model based detection and localization method to deal with abnormal quality levels based on the chlorine measurements and chlorine sensitivity analysis is introduced. The results presented in this work have been obtained in the context of

the i-Sense European Project. The objective of the project is to apply fault detection and isolation techniques for complex systems like CIS.

The structure of the paper is as follows: Section II shows the application of the Barcelona water network considering several selected fault scenarios implemented using a realistic network simulator. Section III presents a FDI model-based method combining both spatial and time series models to perform fault detection and isolation, and Section IV presents a fault detection method which exploits HMM to implement a change detection method in order to perform fault detection. The latter also has fault classification capabilities which are not explored in the current paper, but will be considered in future works. In Section V, fault detection results obtained by each methodology are presented. Finally, conclusions and on-going works are outlined in Section VI.

II. CASE STUDY: BARCELONA WATER TRANSPORT NETWORK

The Barcelona water network, which is taken as a case study in this work, is distributed in 23 different districts covering an area of 424 Km² and providing water to about three million end users. Water managed by the network is obtained from both surface and underground sources, including Ter (surface source) and Llobregat (both surface and underground source) as the most important ones in terms of use and capacity. The water supplied by these sources is distributed through around 4645 Km of pipes to 218 sectors of demand (DMAs) including about 400 control points. Summing up, the complete Barcelona transport network is composed of the following constitutive elements: 3 surface and 7 underground sources, 63 storage tanks, 18 nodes, 79 pumps, 50 valves and 88 demand sectors.

Regarding data management, the Barcelona telecontrol system receives real-time data from 200 control points, which mainly include flow meters (usually installed in the DMA single supply point, so their reading closely fit the actual DMA water demand) and also some pressure sensors.

A. Mathematical model

Similarly as in many physical systems, a full description of the model elements is needed for the real-time monitoring of water networks. The structure of these networks establishes pressure and flow relations between its constitutive elements, e.g., mass conservation at nodes. These relations are given by means of a mathematical model, which should have the following characteristics:

- representative of the hydraulic dynamic response;
- able for real-time operation, i.e. simple enough to allow its computation in a limited period of time;
- balanced in accuracy/complexity trade-off.

In the next section, the hydraulic model used for monitoring the water network is introduced.

1) *Network model description:* The water network model constitutive elements and their basic relationships are introduced in this section. The mass balance expression for the i -th tank is stated as a discrete-time difference equation

$$y_i(k+1) = y_i(k) + \frac{\Delta t}{A_i} (q_{in_i}(k) - q_{out_i}(k)) \quad (1)$$

where $y_i(k)$ is the tank level, A_i is the tank surface, $q_{in_i}(k)$ is the manipulated inflow and $q_{out_i}(k)$ is the outflow, which may include manipulated tank outflow and consumer demands, both given in m³/s.

Moreover, in a water network system nodes are represented as intersections of mains, which mass balance may be expressed as the static equation in (2).

$$\sum_i q_{in_i}(k) = \sum_i q_{out_i}(k) \quad (2)$$

where, similarly to Equation (1), $q_{in_i}(k)$ and $q_{out_i}(k)$ correspond to the inflow and outflow of the i -th subnet node, also given in m³/s.

B. Network Subsystem

The case study presented in this section is called Orioles (Figure 1), which is a subsystem of the whole network that has been chosen due to space restrictions to illustrate the FDI methodology proposed in this work. This part of the network includes the following elements:

- **Tank:** *d175LOR*
- **Actuator with flow sensor:** *iOrioles*
- **Demand with flow sensor:** *c175LOR*
- **Level sensor:** *xd175LOR*

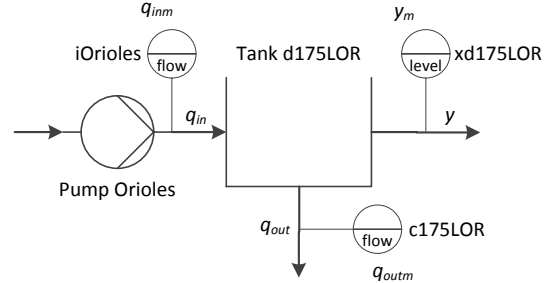


Fig. 1. Orioles subsystem

III. METHOD I: FAULT DIAGNOSIS BASED ON SPATIAL/TIME SERIES MODELS

A. Fault Diagnosis Scheme

The FDI scheme implemented in the Barcelona water network system is shown in Figure 2. In the latter, a model predictive control loop operating on the Barcelona water transport network is presented (Pascual et al., 2011). Moreover, the different elements involved in the FDI method, further detailed in this document, are introduced. These include the spatial consistency (SC) and time-series (TS) models, used to obtain the residuals that are employed by the FDI method in order to detect and isolate the faults appearing in the system. In the next sections, these elements are considered in detail.

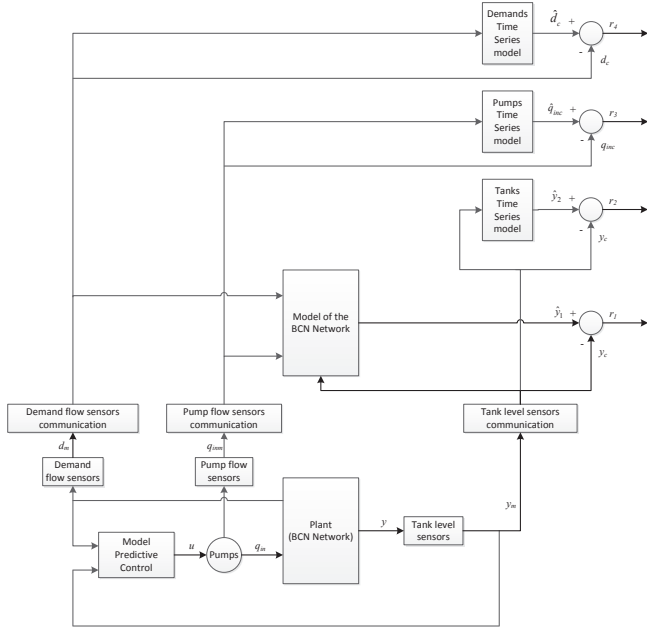


Fig. 2. FDI Scheme for Barcelona Water Network

B. Consistency checks using models

The dataset available for the application of this method, coming from the sensor measurements x , may be divided in three different subsets: training dataset x_a , validation dataset x_b and test dataset x_c . The first one is used for the estimation of the model parameters, the second one is used to validate the model and obtain the corresponding fault detection threshold and the third one is employed for testing. The fault detection module evaluates the nominal residual obtained from the difference between the system measurements and the model prediction, considering the model for the i -th subsystem in input-output regressor form given by

$$r_i(k) = x_i(k) - \hat{x}_i(k) = x_i(k) - \phi_i^T(k)\theta_i \quad (3)$$

where θ_i are the nominal parameters, x_i is the sensor measurement, \hat{x}_i is the model prediction and $\phi_i(k)$ is the regressor vector of dimensions $n_{\theta_i} \times 1$ containing inputs $(u_i(k), u_i(k-1), u_i(k-2), \dots)$ and outputs $(y_i(k), y_i(k-1), y_i(k-2), \dots)$. Considering the uncertainty (e.g. modelling errors, noise) the detection test may be rewritten to check the condition

$$|r_i(k)| < \tau_i \quad (4)$$

where τ_i is the detection threshold and is function of the mean value μ_i and the standard deviation σ_i of the residual (3) for the i -th subsystem model prediction, which is assumed to follow a Gaussian distribution. The threshold in (4) is designed to include the 99.7 % of the values of the whole residual distribution in the faultless situation and hence, it may be

used for fault detection purposes. Test condition (4) can be equivalently expressed as follows:

$$x_i(k) \in [\hat{x}_i(k), \bar{\hat{x}}_i(k)] \quad (5)$$

where $\bar{\hat{x}}_i(k) = \hat{x}_i(k) + \tau_i$ and $\hat{x}_i(k) = \hat{x}_i(k) - \tau_i$, respectively.

C. System residuals

The subsystem Orioles included in the Barcelona water transport network is depicted in Figure 1. The model equations in discrete-time of this subsystems (including the considered faults) are

- Tank (level):

$$y(k+1) = y(k) + \frac{T}{A}[q_{in}(k) - q_{out}(k)] \quad (6)$$

- Pump (flow):

$$q_{in}(k) = q_p(k) + f_p(k) \quad (7)$$

- Tank level sensor:

$$y_m(k) = y(k) + f_{y_m}(k) \quad (8)$$

- Pump flow sensor:

$$q_{in_m}(k) = q_{in}(k) + f_{q_{in_m}}(k) \quad (9)$$

- Demand sector flow sensor:

$$q_{out_m}(k) = q_{out}(k) + f_{q_{out_m}}(k) \quad (10)$$

where $y(k)$ is the actual tank level, $y_m(k)$ is the measured tank level, $q_{out}(k)$ is the actual demand flow, $q_{out_m}(k)$ is the measured demand flow, $q_{in}(k)$ is the actual input tank flow, $q_p(k)$ is the set-point pump flow, $q_{in_m}(k)$ is the measured input flow, $f_c(k)$ is the fault signal related to component c , T is the sampling time and A is the tank surface.

1) *Spatial consistency residuals*: The following ARR (11) may be obtained from Equation (6), using measured variables from the system (Figure 2).

$$y_m(k+1) = y_m(k) + \frac{T}{A}[q_{in_m}(k) - q_{out_m}(k)] \quad (11)$$

Substituting the equations of pumps and sensors (7) to (10) in Equation (11), the faults effect which may be observed in this ARR as follows

$$\begin{aligned} y(k) + f_{y_m}(k) &= y(k-1) + f_{y_m}(k-1) \\ &+ \frac{T}{A}[q_p(k-1) + f_p(k-1) \\ &+ f_{q_{in_m}}(k-1) - q_{out}(k-1) \\ &- f_{q_{out_m}}(k-1)] \end{aligned} \quad (12)$$

Hence, the following residual may be obtained from ARR (11)

$$r_1(k) = y_m(k) - y_m(k-1) + \frac{T}{A}[q_{in_m}(k) - q_{out_m}(k)] \quad (13)$$

2) *Time series residuals*: As may be noted analysing the signature provided by residual $r_1(k)$ (13) in Table I, not all the considered faults are distinguishable without providing extra information apart from this residual (i.e. if just considering $r_1(k)$ all faults have the same signature). Hence, extra residuals may be considered. These residuals can be obtained taking into account that the level in tanks and flow in demand sectors have a daily repetitive behavior that can modelled using a TS model. TS models take advantage of the temporal redundancy of the measured variables. A wide used method for signal forecasting is the Holt Winters (HW) triple exponential smoothing approach ((Winters, 1960), (Makridakis et al., 1998)). This method, which is of wide use because of its simplicity, may be presented in several different versions e.g. additive or damped trend, additive or multiplicative seasonality, single or multiple seasonality. In this work, good performance has been achieved with the additive single seasonality version, which may be implemented as shown next with forecasted value in (14)

$$\hat{x}_{ts}(k+T) = \bar{R}(k) + T\bar{G}(k) + \bar{S}(k-L+T) \quad (14)$$

where \bar{R} is the estimate of the deseasonalized level,

$$\begin{aligned} \bar{R}(k) = & \alpha(x(k) - \bar{S}(k-L)) + (1-\alpha)(\bar{R}(k-1) \\ & + \bar{G}(k-1)) \quad 0 < \alpha < 1 \end{aligned} \quad (15)$$

\bar{G} is the estimate of the trend,

$$\begin{aligned} \bar{G}(k) = & \beta(\bar{R}(k) - \bar{R}(k-1)) \\ & + (1-\beta)\bar{G}(k-1) \quad 0 < \beta < 1 \end{aligned} \quad (16)$$

\bar{S} is the estimate of the seasonal component,

$$\begin{aligned} \bar{S}(k) = & \gamma(x(k) - \bar{R}(k)) \\ & + (1-\gamma)\bar{S}(k-L) \quad 0 < \gamma < 1 \end{aligned} \quad (17)$$

and L is the season periodicity, α , β and γ are the HW parameters (level, trend and season smoothing factors, respectively), x is the measured value and $\hat{x}_{ts}(k+T)$ is the TS model forecasted value for a T step prediction. Hence, analysing the historic records of the measured magnitude in a certain sensor, a HW model is derived and used to validate the current acquired data by this element.

Thus, for each sub-network element with periodic behaviour, a TS HW model can be derived and the following ARR's may be obtained:

- Tank (level) TS:

$$y_m(k) = g(y_m(k-1), \dots, y_m(k-L)) \quad (18)$$

- Demand sector flow TS:

$$q_{out_m}(k) = h(q_{out_m}(k-1), \dots, q_{out_m}(k-L)) \quad (19)$$

- Pump flow TS:

$$q_{in_m}(k) = l(q_{in_m}(k-1), \dots, q_{in_m}(k-L)) \quad (20)$$

where g , h and l are the TS expressions for the tank level sensor, sector demand sensor and pump flow sensor respectively, for data exhibiting a periodicity of L samples.

Replacing the tank level sensor Equation (8), the tank Equation (6) and the pump Equation (7), respectively, in the ARR in (18), the faults effect which may be observed from this ARR are detailed in (21).

$$\begin{aligned} y_m(k) + f_{y_m}(k) = & g(y(k-2) + \\ & \frac{T}{A}[q_p(k-2) + f_p(k-2) - q_{out}(k-2)] \\ & + f_{y_m}(k-1), \dots, \\ & y(k-L-1) + \frac{T}{A}[q_p(k-L-1) \\ & + f_p(k-L-1) - q_{out}(k-L-1)] \\ & + f_{y_m}(k-L)) \end{aligned} \quad (21)$$

Furthermore, replacing the Equations of pump (7), pumps flow sensor (9) and demand flow sensor (10) in ARR's (19) and (20), the faults effect which may be observed from these ARR's are detailed in Equations (22) and (23), respectively.

$$\begin{aligned} q_{out}(k) + f_{q_{out_m}}(k) = & \\ & h(q_{out}(k-1) + f_{q_{out_m}}(k-1), \dots, \\ & q_{out}(k-L) + f_{q_{out_m}}(k-L)) \end{aligned} \quad (22)$$

$$\begin{aligned} q_p(k) + f_p(k) + f_{q_{in_m}}(k) = & \\ & l(q_p(k-1) + f_p(k-1) + f_{q_{in_m}}(k-1), \dots, \\ & q_p(k-L) + f_p(k-L) + f_{q_{in_m}}(k-L)) \end{aligned} \quad (23)$$

Hence, from previous ARR's (18) to (20), the following residuals in Equations (24) to (26) are obtained.

$$r_2(k) = y_m(k) - g(y_m(k-1), \dots, y_m(k-L)) \quad (24)$$

$$r_3(k) = q_{out_m}(k) - h(q_{out_m}(k-1), \dots, q_{out_m}(k-L)) \quad (25)$$

$$r_4(k) = q_{in_m}(k) - h(q_{in_m}(k-1), \dots, q_{in_m}(k-L)) \quad (26)$$

Notice that in residual $r_2(k)$ in (24), there are hidden dependencies of previous level measurements with certain faults. For example, considering Equation (21) coming from the ARR in (18) and the tank Equation in (6), it can be noticed that the fault in the sensor measuring the demand flow serving this tank (i.e. $f_{q_{out_m}}$) cannot be detected with residual $r_2(k)$. The reason of this hidden dependency is because this measured flow is not used in Equation (21), but determined by the real demand flow. Furthermore, also notice that TS residuals are only useful for diagnosis of faults related to elements with periodic behavior. This is applicable to all the faulty elements considered in this work and also to the major part of the elements of the network, but still there are some which are not expected to evolve periodically (e.g. pump stations related to a non-regular demand node with random demand pattern). Hence, in order to apply this fault isolation methodology, a previous study of the nature of the elements

TABLE I
FAULTS SIGNATURES

	f_{y_m}	$f_{q_{out_m}}$	$f_{q_{in_m}}$	f_p
r_1	1	1	1	1
r_2	1	0	0	1
r_3	0	1	0	0
r_4	0	0	1	1

involved must be performed. As a counterpart, these residuals may be computed with information provided from a single sensor, which may be advantageous in certain applications when more information is not available, e.g. intelligent sensors with embedded diagnosis unit (Alippi et al. (2012)). Hence, in this work spatial consistency residuals will be used for fault detection, whilst TS residuals will be employed for fault isolation purposes.

From residuals in Equations (13), (24), (25) and (26), the theoretical binary fault signature matrix in Table (I) may be presented.

IV. METHOD II: FAULT DIAGNOSIS BASED ON HIDDEN MARKOV MODELS CHANGE DETECTION METHOD

The HMM-based Change Detection Test (CDT) assumes that the relationship between two generic correlated data streams y_m and q_{in_m} , q_{in_m} used to infer y_m , can be described through a linear input-output dynamic model of the form

$$y_m(k) = f_h(y_m(k-1), y_m(k-2), \dots, y_m(k-k_i), q_{in_m}(k), q_{in_m}(k-1), \dots, q_{in_m}(k-k_j))$$

where f is a function of linear time-invariant type (e.g. ARX, ARMAX, OE) in its parameters h and k_i and k_j are the orders of the model.

Let $O_{i,T_0} = \{y_m(t), t = 1, \dots, T_0\}$ and $O_{j,T_0} = \{q_{in_m}(t), t = 1, \dots, T_0\}$ be the data sequence of the i -th and j -th sensors which are used to extract the linear time invariant model, the parameters of which are used to train the HMM. In general, HMMs constitute probabilistic machines able to automatically identify a sequence of patterns within a stream of data provided that enough data to train the model are given (Rabiner, 1989). A HMM is a probabilistic oriented graph composed of:

- a number of states, S ,
- the probability density function associated with each state modelled as a mixture of Gaussians (GMM), $P(x|h) = \sum_{k=1}^K p_k p(x|h_{(k)})$, where p_k s are the mixture weights, x is a continuous-valued data vector (e.g. measurements or features), $h_{(k)}$ represents the k -th component of the vector, $h = [\sum, \mu]$, $p(x|h_{(k)}) = \frac{1}{(2\pi)^{d/2} |\sum_k|} e^{-\frac{1}{2}(x-\mu_k)^t \sum_k^{-1}(x-\mu_k)}$ the state transition probability matrix $A = \{a_{ij}\}$ where entry a_{ij} represents the probability of moving from state j at time t to state i at time $t+1$.

- the initial state distribution $\pi = \{\bar{\pi}_i\}$, where $\bar{\pi}_i$ corresponds to the probability that the HMM starts in state i .

A HMM is initially built by exploiting a faulty-free training sequence. During operational time, novel data are then compared with the structure present in the HMM by means of a log-likelihood measure. To this end the HMM defined by $S, P(x|h), A$ and π can generate the most likely sequence of states that resulted at the observed data stream.

When the incoming data stream is fed to the HMM, it produces a log-likelihood measuring how similar is the probabilistic pattern of observed data to the training data. When the log-likelihood decreases below a predefined threshold, a change is detected. In principle, this change may correspond either to a fault (sensor malfunction, network error etc.) or a change in the sensed variable (change in the surrounding environment). Depending on these cases, one can infer useful information and take appropriate actions, e.g. repair the faulty element or notify the appropriate personnel regarding an emergency situation (Alippi et al., 2013).

The probability of the observation sequence $O = O_1 O_2 \dots O_T$ given the model $H = \{S, P(x|h), A, \pi\}$, i.e., $P(O|H)$ is computed by listing every possible state sequence of length equal to the length of the observations T . Let $Q = q_1 q_2 \dots q_T$ be an instance of all the possible realizations of sequences of states of length T . The probability that the observation sequence O has been generated by H for the generic sequence Q is $P(O|Q, H) = \prod_{t=1}^T P(O_t|q_t, H)$.

Since we have assumed statistical independence of observations, $P(O|Q, H) = \prod_{t=1}^T P(O_t|q_t, H)$. This probability can be written as $P(Q|H) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{T-1} q_T}$. Now, the joint probability of O and a given Q is the product of the previous two terms, i.e., $P(O, Q|H) = P(O|Q, H) P(Q|H)$. Thus the probability of O given the model H is obtained by summing this joint probability over all possible state sequences Q giving:

$$P(O|H) = \sum_{\forall Q} P(O|Q, H) P(Q|H)$$

A. The HMM-CDT algorithm

We summarize the HMM based change detection algorithm in Algorithm 1.

The training set corresponds to $O_{i,T_0}, 1 \leq i \leq N$, selected in the initial data stream so that no changes are there present. The particular model coefficients are used to train the HMM which is characterized by $\{S, P(x|h), A, \pi\}$. The obtained HMM is then considered as a representation of the change-free data class. Subsequently, it is used to estimate the threshold T_h on a validation set, denoted as $O_{i,T_v}, N+1 \leq i \leq Z$. When unknown data are analysed they are first windowed and the model coefficients with respect to each window are computed and inserted into the trained HMM. The log-likelihood associated with each window is then calculated and if below T_h it is

1. Build the change-free class HMM,
 $H_{change-free} = \{S, P, A, \pi\}$ from the vectors of parameters $h_1 \dots h_d$ each of which associated with a linear dynamic model applied to the training data $O_{i,T_0, i=1, \dots, d}$ windowed using length M overlapping by $M - 1$;
2. Determine threshold T_h as the minimum log-likelihood value of h_{T_0} on O_{T_v} ;
3. Window the incoming novel data as above, which results in windows $W = W_1 \dots W_n$;

repeat

4. $j=1$;
5. Compute the parameter vectors of the $j - th$ dynamic model h_j with respect to W_j ;
6. Compute the log-likelihood
 $L_{W_j} = P(h_1 \dots h_j | H_{change-free}) = P(O|Q, H) = \prod_{i=1}^T P(O_i | q_i, H_{change-free})$;
7. **if** $L_{W_j} < T_h$ **then**
| W_j contains data associated with a change, alarm notification
| **else**
| W_j contains data coming from the normal working modality
| **end**
8. $j = j + 1$;

until (1);

Algorithm 1: The HMM-based Change Detection Algorithm

considered to contain data associated with a change. When the opposite holds, the data are considered to be change-free.

In the case of the Orioles subsystem of the Barcelona water distribution network, the relationship pattern of the measured tank level and the measured input flow data streams is modelled. The motivation behind modelling the specific relationship is that $y_m(k)$ and $q_{in_m}(k)$ comprise the natural input/output variables of the Orioles system (Figure 1). It should be noted that the measured demand follows a pattern imposed by the consumers and may be considered as a disturbance factor. In faulty conditions, the considered relationship will exhibit changes which depend on the induced fault, thus its monitoring is useful for fault detection purposes.

V. RESULTS

A. Fault Scenarios

The simulator used to exhaustively test the proposed methods before its application to the real Barcelona water network has been developed in MATLAB/SIMULINK environment, using a model calibrated and validated with real data providing a good degree of representativeness of the actual network. As introduced in Section II-B, the application of the purposed methodologies is focused on iOrioles network subsystem (Figure 1) for illustrative purposes.

One of the features of the Barcelona water network simulator is the possibility of setting faults of different kinds in

distinct elements of this network. The faults considered are introduced in the actuators and sensors, and may be of the types detailed next:

- **Freezing:** Measurements from sensors are normally time varying functions, but when this fault is occurring the signal remains constant after a given fault initial time
- **Offset:** When this fault is taking place, a constant magnitude is added to the sensor measurements after a given fault initial time
- **Drift:** When this fault occurs, a ramp function is added to the sensor measurements after a given fault initial time

To test and adjust the methods presented in this paper, different fault scenarios have been defined, including random normally distributed measurement noise of ± 1 % full scale. The faults considered, $f(t) = f_o \beta(t)$, are both of abrupt or incipient nature, as defined as follows

- Abrupt fault:

$$\beta(t) = \begin{cases} 0, & t < t_f \\ 1, & t \geq t_f \end{cases} \quad (27)$$

- Incipient fault:

$$\beta(t) = \begin{cases} 0, & t < t_f \\ 1 - e^{-\rho(t-t_f)}, & t \geq t_f \end{cases} \quad (28)$$

where $\rho > 0$ is the constant describing the evolution rate of the fault and t_f is the time instant when the fault is occurring. The parametrization of these faults is depicted in Table II.

It should be mentioned that the HMM-based CDT uses linear models of ARX type for the extraction of the parameters h .

The dataset considered to implement these faults scenarios lasts for seven months, with a sampling period $T = 1$ hour and with a fault appearing at $t_f = 744 \cdot T$ in iOrioles pump sensor ($f_{q_{in_m}}$), installed in Orioles subsystem (Figure 1). Regarding methods initialization, the first thirteen days of data are used as training dataset to identify the model parameters (O_{T_0}), the next thirteen days (O_{T_v}) are used as validation dataset to obtain the corresponding fault detection threshold and the rest of the data is used as test dataset. The numerical results are presented by using different figures of merit

- **Delay:** Number of samples needed by the fault detection method to detect a certain fault
- **False Positives (FP):** Percentage of test dataset faultless samples (i.e. non-affected by a certain fault) that are determined as faulty by the fault detection method
- **False Negatives (FN):** Percentage of test dataset faulty samples (i.e. affected by a certain fault) that are determined as faultless by the fault detection method

In Table II, the fault detection results achieved by both methods are presented. Except for some low magnitude faults (Id.1 and Id.5) and freezing faults (Id.17 and Id.18) where high FN rates are obtained by Method I, or also for some low magnitude faults where high delay rates are obtained by Method I and Method II, it may be observed how both methodologies perform in a good manner and obtain better

TABLE II
FAULTS PARAMETRIZATION AND DETECTION RESULTS (MFD STANDS FOR MAXIMUM FLOW/DEMAND)

Fault Id.	Type of fault	Magnitude	Temporal/Spatial model-based method			HMM based method		
			Delay [# of samples]	FP [%]	FN [%]	Delay [# of samples]	FP [%]	FN [%]
1	Offset abrupt	5 % MFD	4	0	59.13	14	0	4.1
2	Offset abrupt	10 % MFD	2	0	2.18	12	0	1.2
3	Offset abrupt	15 % MFD	2	0	0.05	11	0	0.3
4	Offset abrupt	25 % MFD	2	0	0.05	9	0	0
5	Offset incipient	5 % MFD	25	0	59.39	15	0	3.5
6	Offset incipient	10 % MFD	12	0	2.83	12	0	1.9
7	Offset incipient	15 % MFD	9	0	0.37	11	0	0.2
8	Offset incipient	25 % MFD	9	0	0.23	9	0	0
9	Drift abrupt	0.1 % MFD	36	0	1.33	16	0	1.1
10	Drift abrupt	1 % MFD	9	0	0.21	12	0	0.5
11	Drift abrupt	5 % MFD	4	0	0.09	8	0	0.3
12	Drift abrupt	10 % MFD	3	0	0.07	7	0	0
13	Drift incipient	0.1 % MFD	50	0	1.52	22	0	0.3
14	Drift incipient	1 % MFD	12	0	0.35	14	0	0.2
15	Drift incipient	5 % MFD	9	0	0.21	14	0	0.2
16	Drift incipient	10 % MFD	7	0	0.1611	12	0	0
17	Freezing abrupt	-	7	0	40.52	11	0	3.7
18	Freezing incipient	-	19	0	40.82	14	0	4.2

results as the magnitude of each fault is increasing. From results in Table II, it may be also noted how performances attained by both methods are complementary in most of the situations, since generally Method II seems to behave better for low magnitude and freezing faults whilst Method I obtains faster detection rates. This may motivate further work to be done on the potential combination of both approaches in order to improve the overall diagnosis performance. A possible approach could be to use Method I and Method II together to implement fault detection, trying to take advantage of the benefits of each method separately. Also, further works will be performed to apply the isolation potential of both methods and extend the fault detection comparison results presented in this work to the isolation and classification performance comparison of both methods.

VI. CONCLUSIONS

In this work, a comparison between two distinct fault diagnosis frameworks has been developed, that are respectively based on a temporal/spatial model-based fault diagnosis method and a hidden Markov model change detection method. Comparison when used for fault detection has been performed using a case study based on a benchmark of the Barcelona water transport network. The performance of both methods has been measured using different figures of merit, showing good fault detection performance for distinct faulty scenarios affecting a part of this network. The methods considered also have isolation and classification capabilities, as has been already introduced for the temporal/spatial model-based fault diagnosis method, which will be compared in the next steps of this work. Future work may also be performed on the simultaneous usage of both methods in order to improve fault detection and isolation performance achieved by each method separately,

considering the advantages of each approach independently, since each one addresses the problem from a different point of view due to their heterogeneous nature.

ACKNOWLEDGEMENT

This work was supported by the FP7 EU project i-Sense, Making Sense of Nonsense, Contract No: INSFO-ICT-270428.

REFERENCES

- Alippi, C., Ntalampiras, S., Roveri, M., June 2012. An hmm-based change detection method for intelligent embedded sensors. In: Neural Networks (IJCNN), The 2012 International Joint Conference on. pp. 1–7.
- Alippi, C., Ntalampiras, S., Roveri, M., 2013. A cognitive fault diagnosis system for sensor networks. IEEE Transactions on Neural Networks and Learning Systems, conditionally accepted.
- Makridakis, S., Wheelwright, S., Hyndman, R., 1998. Forecasting methods and applications. John Wiley & Sons.
- Nejjari, F., Pérez, R., Puig, V., Quevedo, J., Sarrate, R., Cugueró, M. A., Sanz, G., Mirats, J. M., 2012. Abnormal quality detection and isolation in water distribution networks using simulation models. Drinking Water Engineering and Science 5 (1), 67–72.
URL <http://www.drink-water-eng-sci.net/5/67/2012/>
- Pascual, J., Romera, J., Puig, V., Creus, R., Minoves, M., 2011. Operational predictive optimal control of barcelona water transport network. In: World Congress of the International Federation of Automatic Control. "Proceedings of the 18th IFAC World Congress". Milan.
- Quevedo, J., Cugueró, M. A., Pérez, R., Nejjari, F., Puig, V., Mirats, J. M., 2011. Leakage location in water distribution networks based on correlation measurement of pressure sensors. In: 8th IWA Symposium on System Analysis and Integrated Assessment (WATERMATEX 2011). San Sebastian.
- Rabiner, L. R., 1989. A tutorial on hidden markov models and selected applications in speech recognition. Proceedings of the IEEE, 257–286.
- Winters, P. R., 1960. Forecasting sales by exponentially weighted moving averages. Management Science 6 (52), 324–342.
URL <http://www.jstor.org/stable/10.2307/2627346>