united nations educational, scientific and cultural organization

international atomic energy agency

the
abdus salam
international centre for theoretical physics

SMR1327/32

Summer School on Mathematical Control Theory

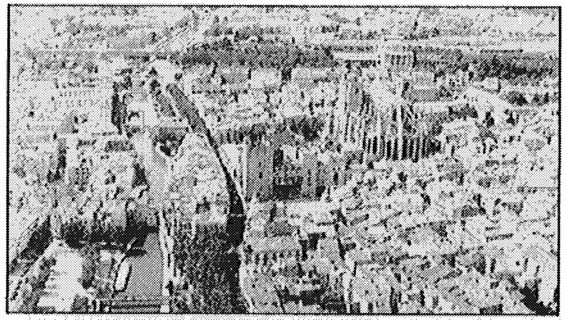
(3 - 28 September 2001)

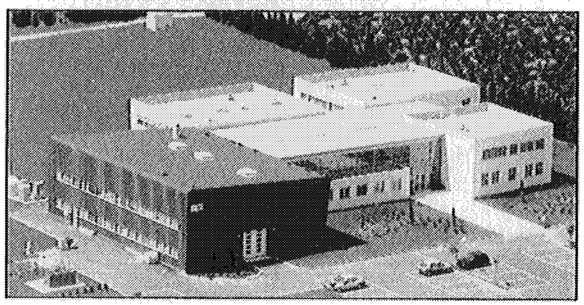
Experimental illustrations of on-line diagnosis of wastewater treatment processes using analytical and heuristic approaches

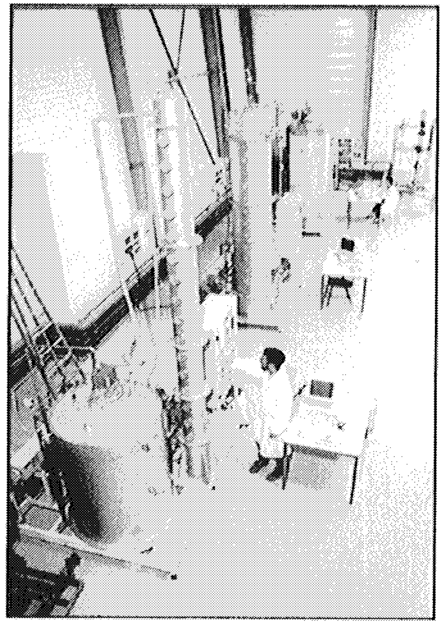
Jean-Philippe Steyer
Equipe de Recherche en Automatique et Procédés
Laboratoire de Biotechnologie de l'Environnement
INRA
Avenue des Etangs
11100 Narbonne
France

Experimental illustrations of on-line diagnosis of wastewater treatment processes using analytical and heuristic approaches

Dr. Jean-Philippe Steyer
PEACE (Process Engineering And Control Engineering) Research Group
Laboratoire de Biotechnologie de l'Environnement - INRA
Narbonne - France
email : steyer@ensam.inra.fr - http://www.ensam.inra.fr/narbonne







INRA

Contents of the Presentation

1. ***Some reasons for diagnosis***

2. ***Diagnosis using analytical models***

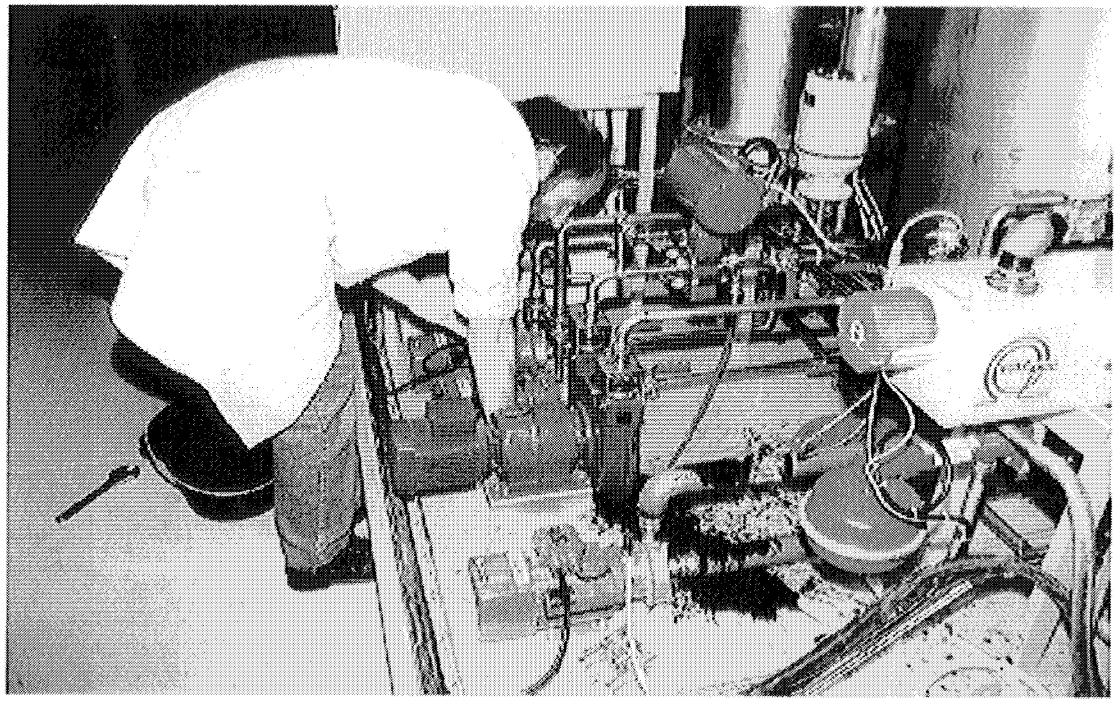
3. ***Diagnosis using heuristic knowledge***

4. ***Conclusions and perspectives***

Process Control is not only Optimality Calculation

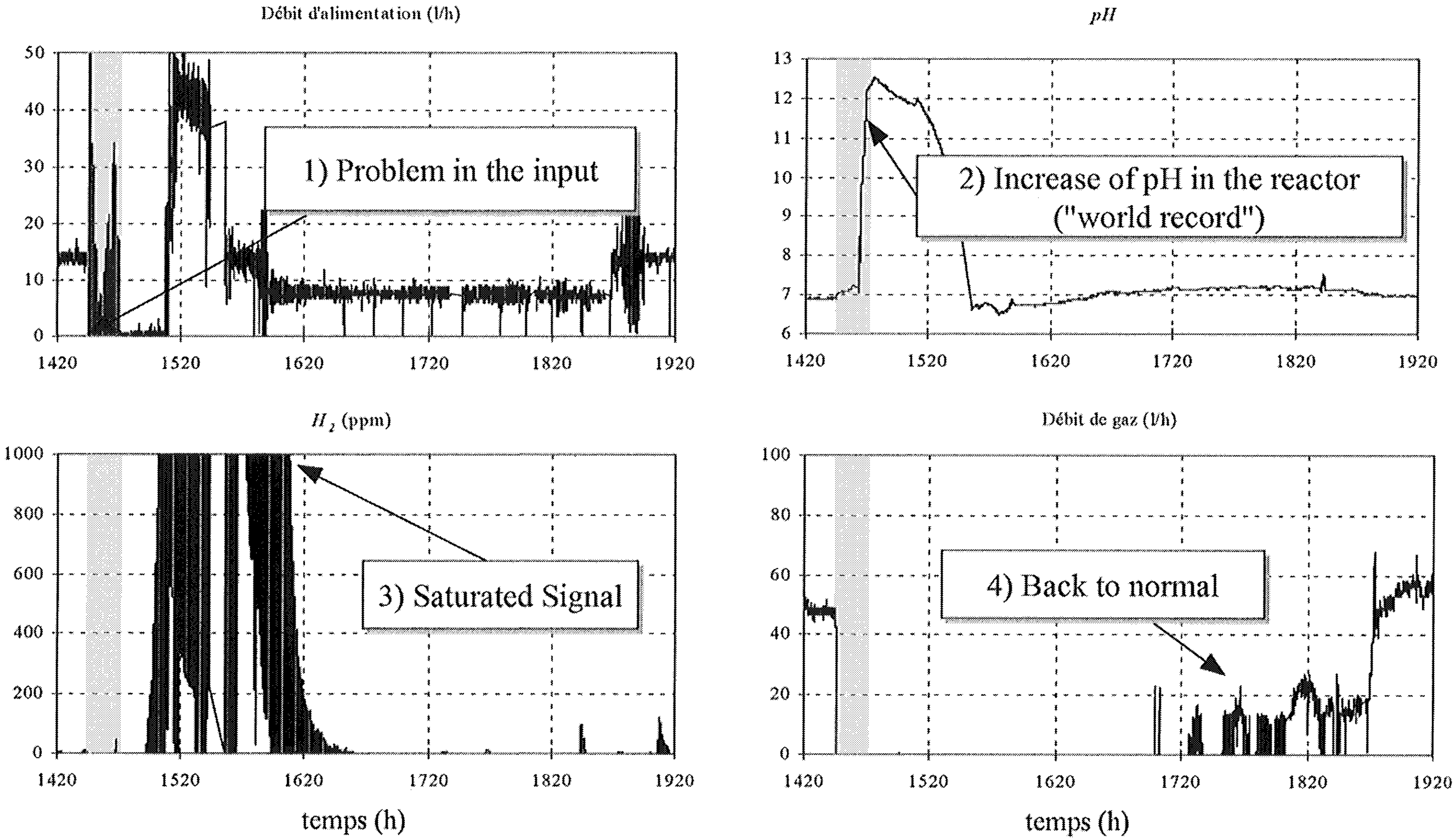
The prime objective of a control law is
to help the human operator and to relieve him from the tedious tasks
(that are sometimes not of the most interest) ...

However, technical problems can lead to complete opposite situations !



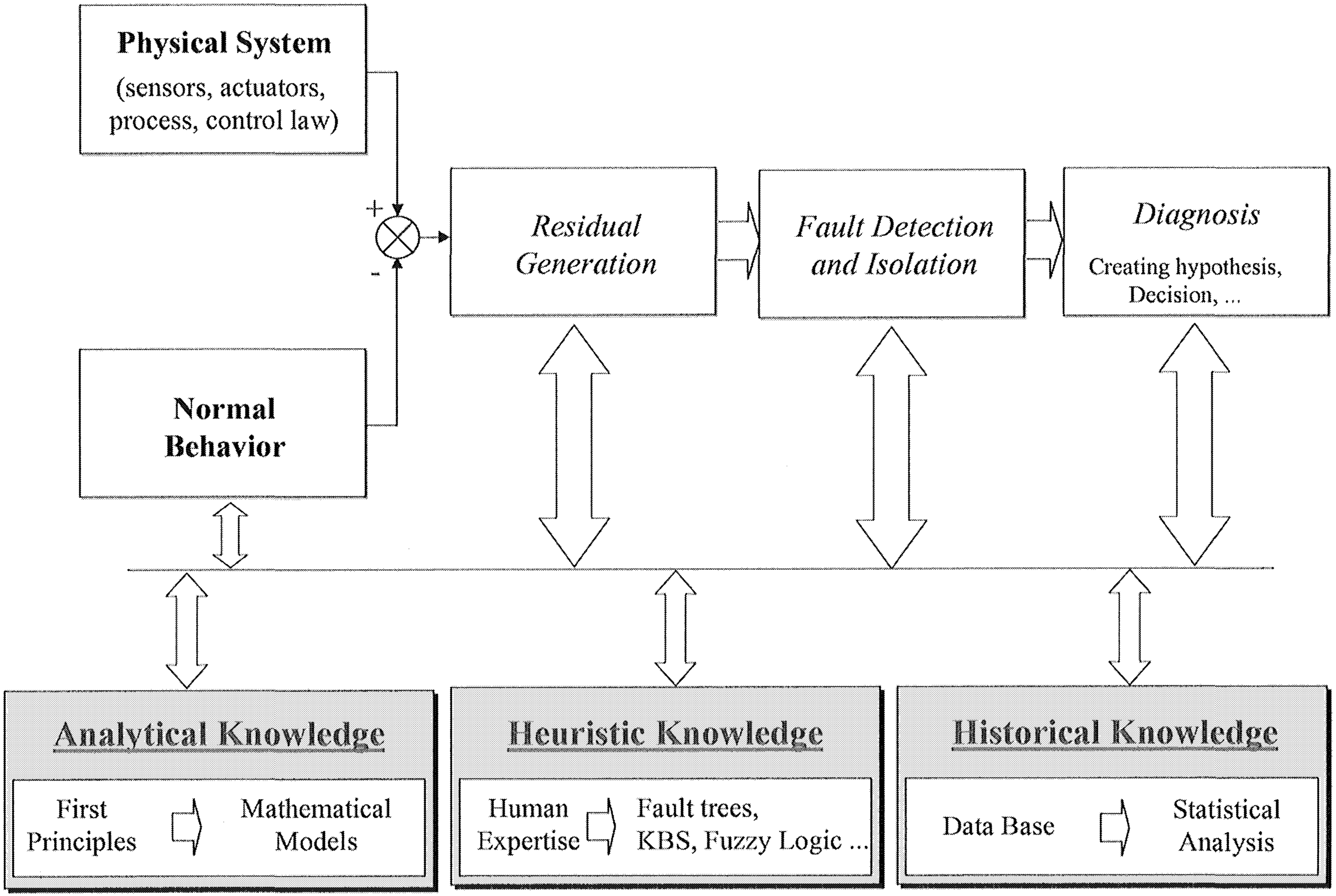
Example of the consequences of a pipe clogging : the adjustable spanner is then the only "solution"

Reasons for Diagnosis



Complete stop of the process for one month and new start-up (with new sludges) required

Possible Approaches for Diagnosis



Contents of the Presentation

Diagnosis using Analytical Models

1) An Off-line Application

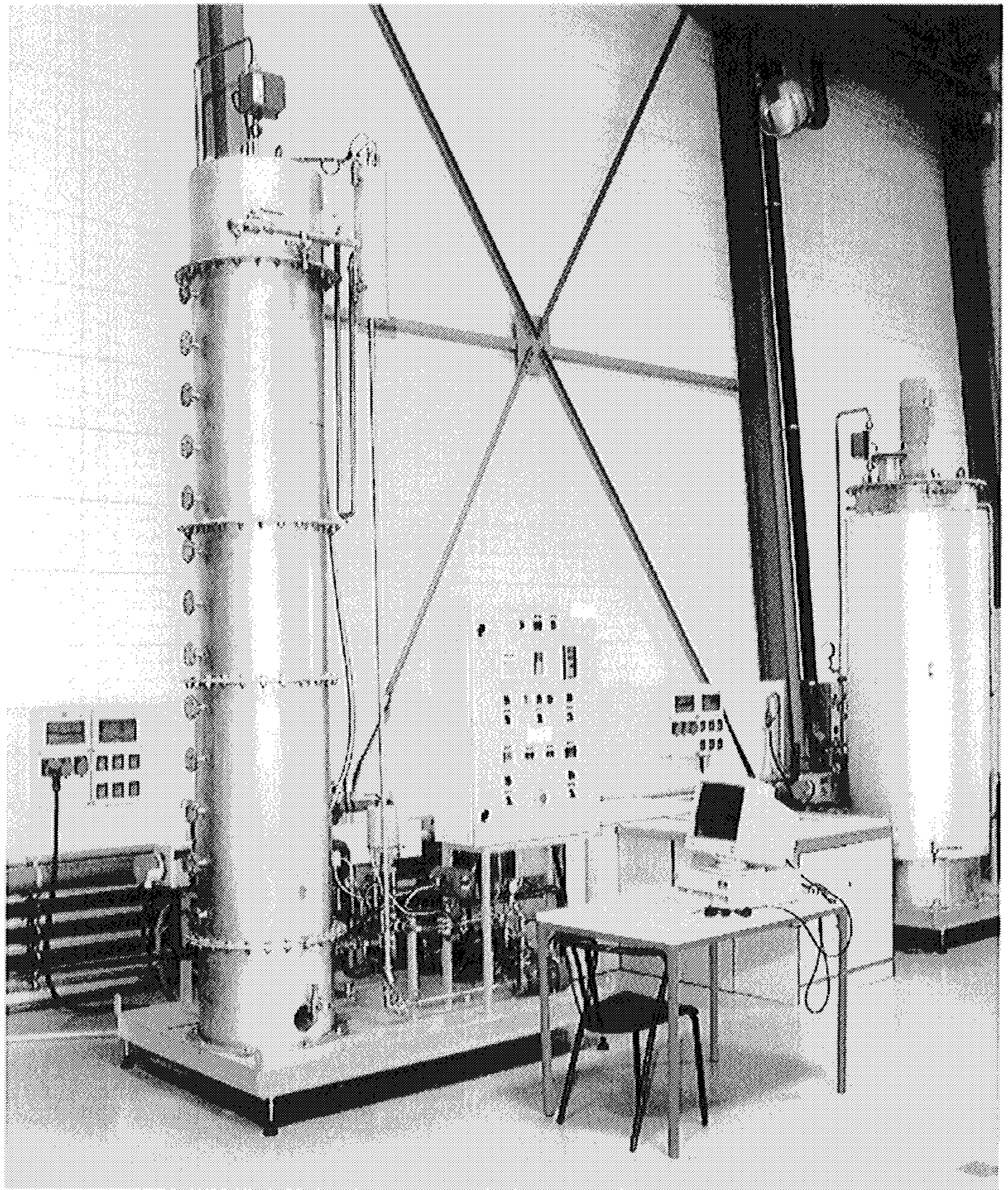
The process used in Narbonne

Influent : **Raw industrial distillery vinasses**

Reactor : **Circular column**
Up-flow fixed bed reactor
- 3.5 m height,
- 0.6 m diameter,
- 982 liters of total volume.

Media : **Cloisonyl**
- Specific surf. : 180 m^2/m^3
- Volume : 33.7 liters

Total effective volume : **948 liters**



The "AMOCO" Anaerobic Digestion Model (EU FAIR Project CT 1198)

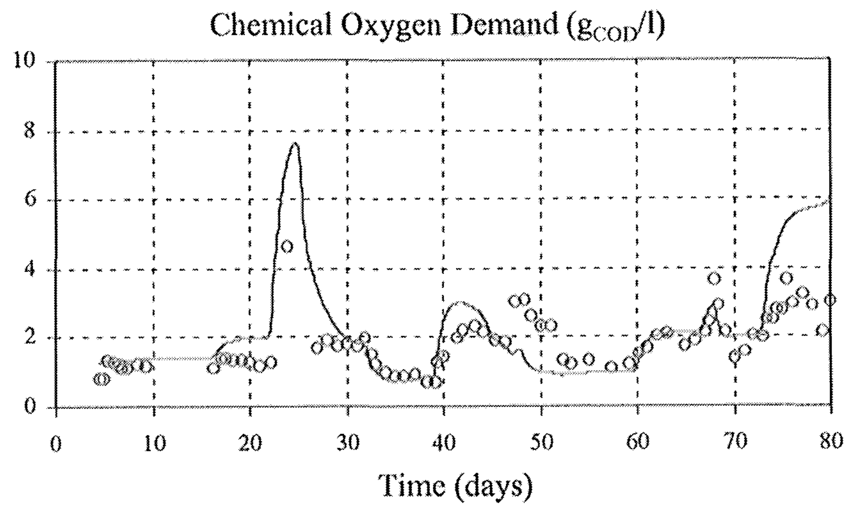
From Mass Balance

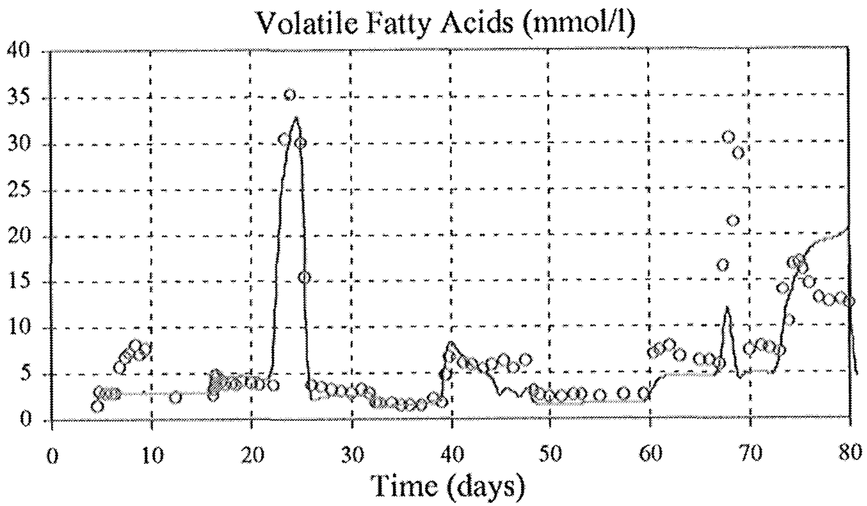
$$
\begin{cases}
\dot{X}_1 = (\mu_1 - \alpha D) X_1 \\
\dot{X}_2 = (\mu_2 - \alpha D) X_2 \\
\dot{Z} = D(Z^i - Z) \\
\dot{S}_1 = D(S_1^i - S_1) - k_1 \mu_1 X_1 \\
\dot{S}_2 = D(S_2^i - S_2) + k_2 \mu_1 X_1 - k_3 \mu_2 X_2 \\
\dot{C}_{TI} = D(C_{TI}^i - C_{TI}) + k_7 (k_8 P_{CO_2} + Z - C_{TI} - S_2) + k_4 \mu_1 X_1 + k_5 \mu_2 X_2
\end{cases}
$$

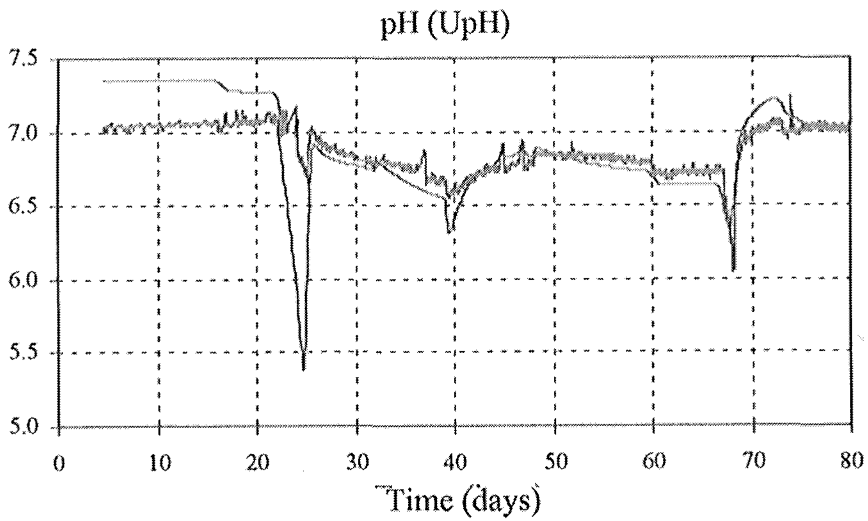
With $\mu_1 = \mu_{\max 1} \dfrac{S_1}{K_{S_1} + S_1}$ (*i.e.*, limitation by organic matter)

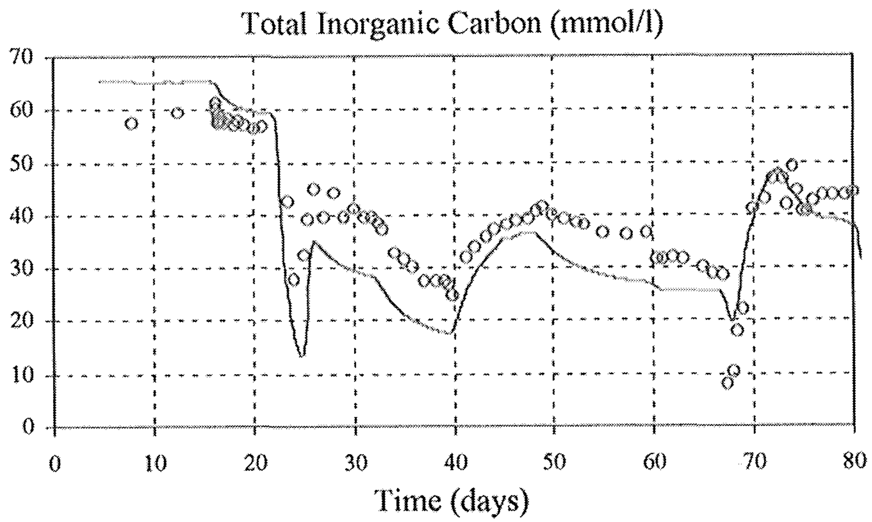
and $\mu_2 = \mu_{\max 2} \dfrac{S_2}{K_{S_2} + S_2 + \left(\dfrac{S_2}{K_{I_2}}\right)^2}$ (*i.e.*, limitation and inhibition by VFA)

Dynamic simulations vs. experimental data (1997)









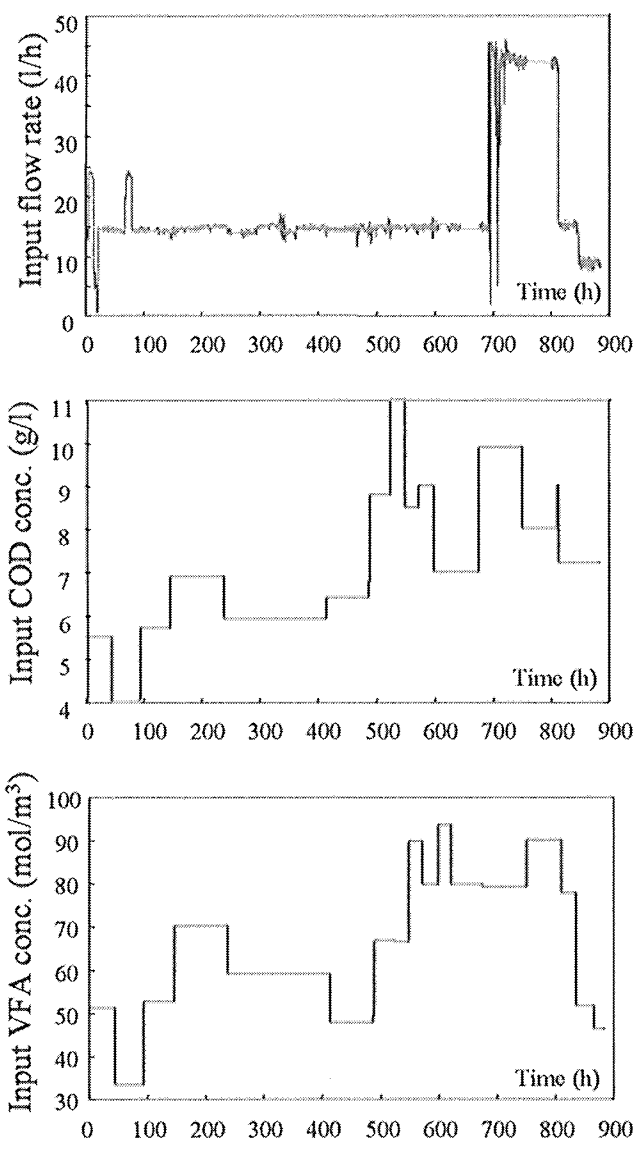
Dynamical Model of the Anaerobic Digestion

• From the initial model developped in 1997, 6 distinct sets of parameters have been identified. Their range of variation is represented in the following Table :

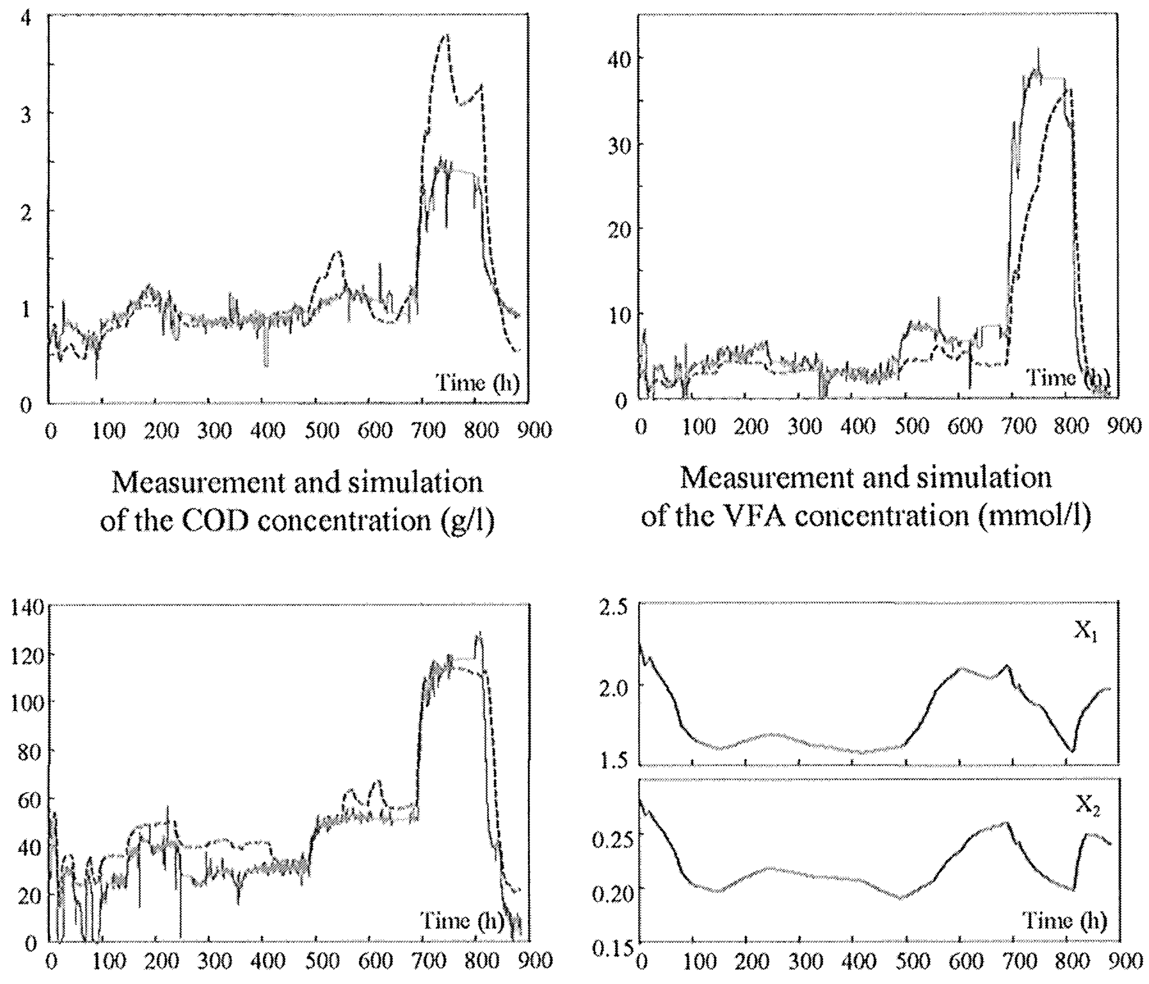
Name	Meaning	Min	Max	Unit
k_1	Yield coefficient for COD degradation	6.60	86.11	g COD/g X_1
k_2	Yield coefficient for fatty acid production	7.80	181.2	mmol VFA/g X_1
k_3	Yield coefficient for fatty acid consumption	268	1814	mmol VFA/g X_2
k_4	Yield coefficient for CO_2 production due to X_1	12.4	230	mmol CO_2/g X_1
k_5	Yield coefficient for CO_2 production due to X_2	273	1924	mmol CO_2/g X_2
k_6	Yield coefficient for CH_4 production	315	2696	mmol CH_4/g X
k_7	Liquid/gas transfer rate	19.8	500	day^{-1}
k_8	Henry's constant	16.0	26.7	mmol CO_2/l-atm
α	Proportion of dilution rate for bacteria	0.30	0.50	
μ_{max1}	Maximum acidogenic biomass growth rate	1.20	1.40	day^{-1}
μ_{max2}	Maximum methanogenic biomass growth rate	0.50	0.85	day^{-1}
K_{S1}	Saturation parameter associated with S_1	3.72	10.7	g COD/l
K_{S2}	Saturation parameter associated with S_2	5.28	18.0	mmol VFA/l
K_{I2}	Inhibition constant associated with S_2	16.0	25.0	mmol VFA/l

Two years later (in February 1999)

Input of the process



Output of the process



Measurement and simulation of the COD concentration (g/l)

Measurement and simulation of the VFA concentration (mmol/l)

Measurement and simulation of the methane flow rate (l/h)

Simulation of the X_1 and X_2 concentrations in the reactor (g/l)

Three years later (in February 2000)

- In February 2000, no parameter set was found to correctly match the proces data. Thus, ***using the parameters set from 1999,*** we decided to add the initial biomass concentrations in the optimization criterion :

$$\begin{cases} J = \min\limits_{\theta} \dfrac{1}{n}\left(\sum\limits_{i=1}^{n} \left\| S_1^{norm} - \hat{S}_1^{norm} \right\|_2 + \sum\limits_{i=1}^{n} \left\| S_2^{norm} - \hat{S}_2^{norm} \right\|_2 + \sum\limits_{i=1}^{n} \left\| Q(CH_4)_{norm} - \hat{Q}(CH_4)_{norm} \right\|_2 \right) \\ \text{subject to } \; 0 < \theta_{\mathrm{i}} \end{cases}$$

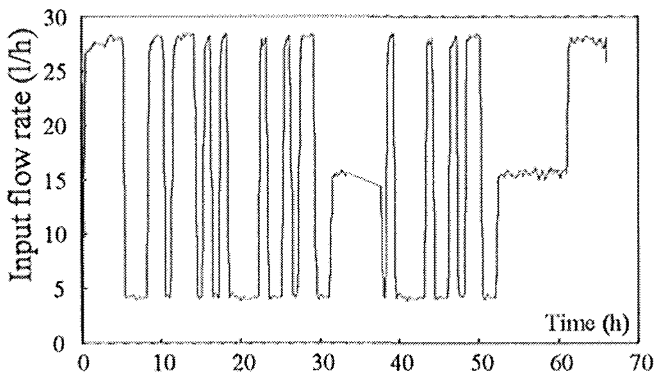
with $\theta_1 = X_1(0)$ and $\theta_2 = X_2(0)$.

Best simulation results obtained in February 2000

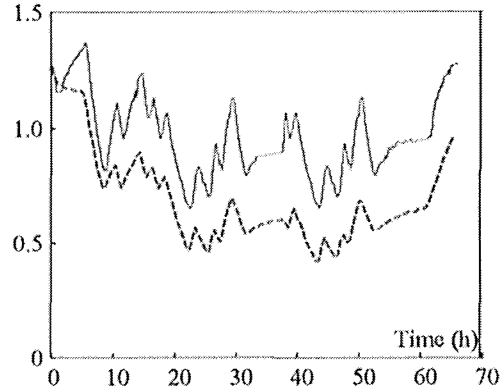
(using the same working volume than in 1997)

Input of the process

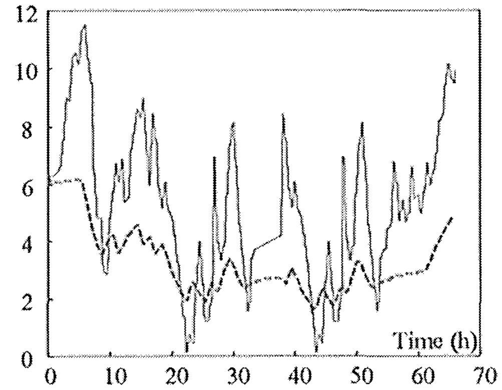
Constant input concentration but changing input flow rate



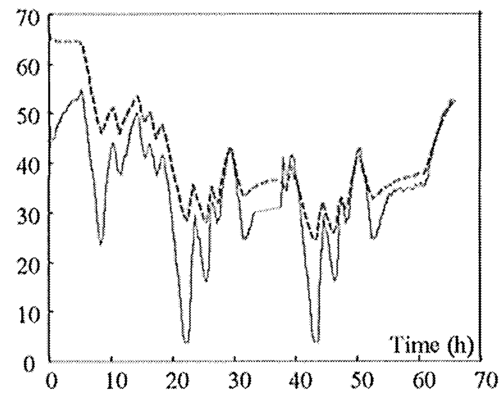
Output of the process



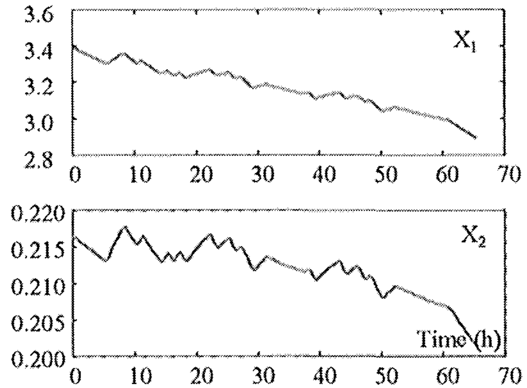
Measurement and simulation of the COD concentration (g/l)



Measurement and simulation of the VFA concentration (mmol/l)



Measurement and simulation of the methane flow rate (l/h)



Simulation of the X_1 and X_2 concentrations in the reactor (g/l)

$J(\theta)=1.0463$ with $X_1(0) = 3.38$ kg/m³ and $X_2(0) = 0.22$ kg/m³

February 2000 : Is there a clogging of the process ?

- A clogging of the process is suspected.

- Both the initial biomass concentrations ***AND*** the active volume are then re-identified in minimizing the following criterion (still using other parameters determined from 1999 data) :

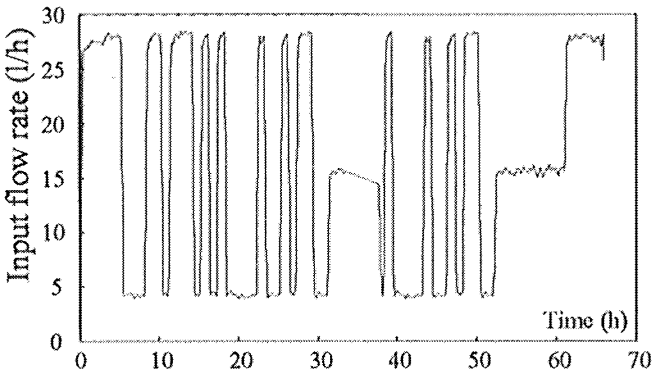
$$\begin{cases} J = \min\limits_{\theta} \frac{1}{n}\left(\sum_{i=1}^{n} \left\| S_1^{norm} - \hat{S}_1^{norm} \right\|_2 + \sum_{i=1}^{n} \left\| S_2^{norm} - \hat{S}_2^{norm} \right\|_2 + \sum_{i=1}^{n} \left\| Q(CH_4)_{norm} - \hat{Q}(CH_4)_{norm} \right\|_2 \right) \\ \text{subject to } \; 0 < \theta_\text{i} \end{cases}$$

with $\theta_1 = X_1(0), \;\; \theta_2 = X_2(0) \;\; and \;\; \theta_3 = V$

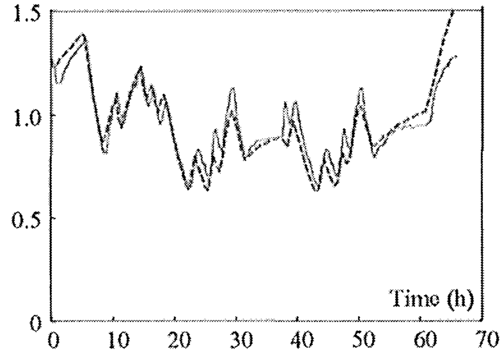
After optimization of the initial biomass concentrations *and* of the active volume

Input of the process

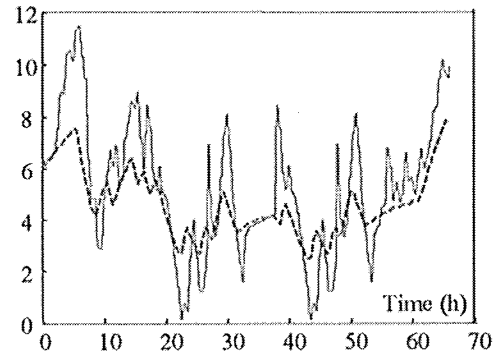
Constant input concentration but changing input flow rate



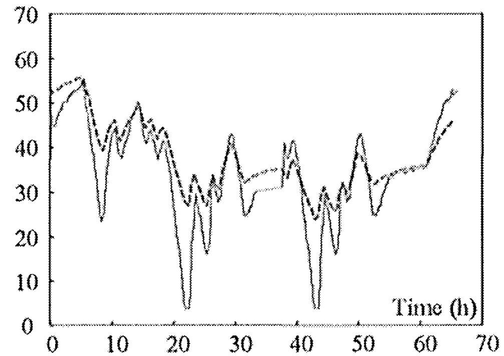
Output of the process



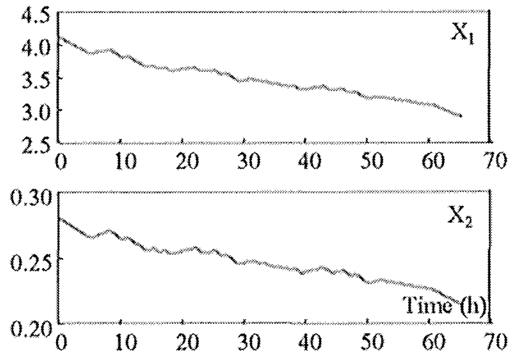
Measurement and simulation of the COD concentration (g/l)



Measurement and simulation of the VFA concentration (mmol/l)



Measurement and simulation of the methane flow rate (l/h)



Simulation of the X_1 and X_2 concentrations in the reactor (g/l)

Before : $J(\theta)=1.0463$ with $X_1(0) = 3.38$ kg/m³ $X_2(0) = 0.22$ kg/m³ and $V_{opt} = 948$ liters

Now : $J(\theta)=0.7636$ with $X_1(0) = 4.11$ kg/m³, $X_2(0) = 0.28$ kg/m³ and $V_{opt} = 595$ liters

INRA

Experimental determination of the HRT

- In order to verify our hypothesis (*i.e.*, a clogging of the process), the experimental determination of the HRT is realized.

- A hydrodynamic study is carried out by means of tracer pulse experiments (using lithium chloride) in March 2000. Two experiments were done (before and after the washing).

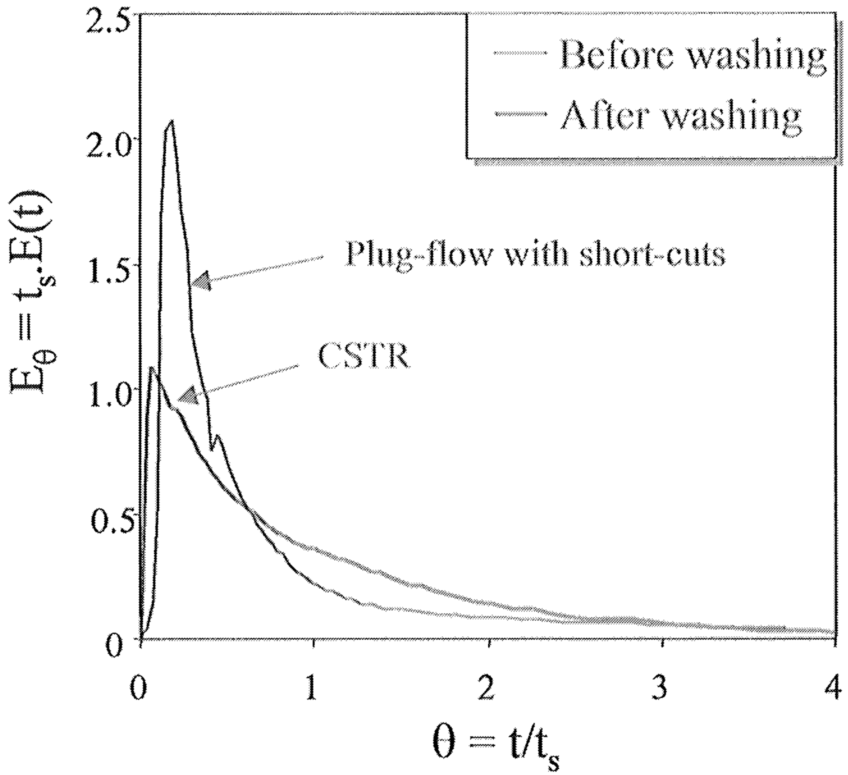
- A classical mathematical treatment was applied to the results. The Residence Time Distribution function E(t) was estimated from the Li^+ concentration as :

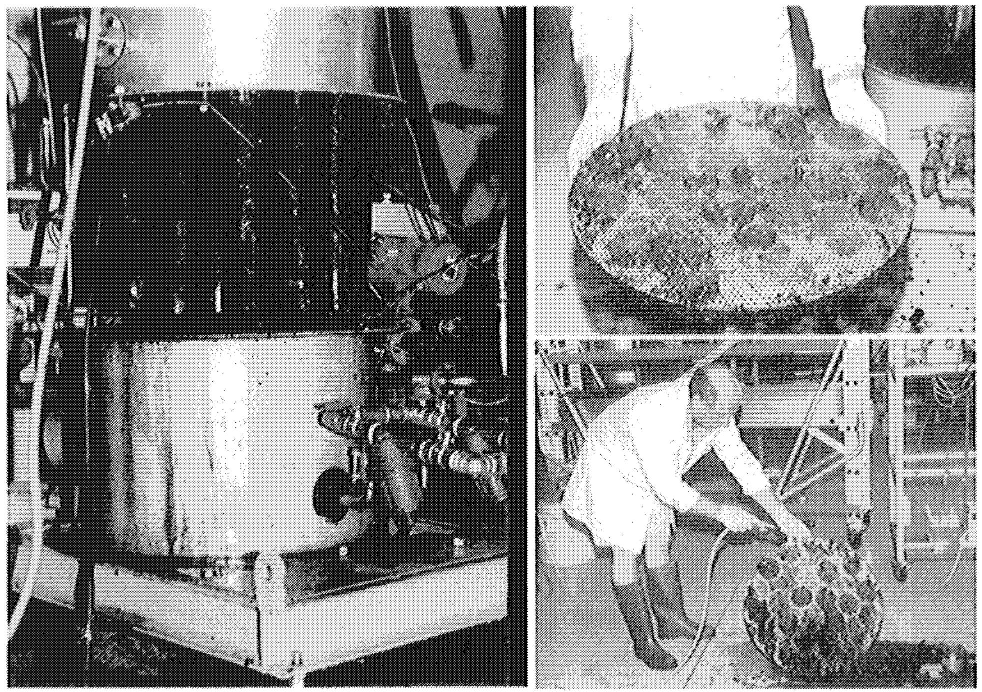
$$E(t_j) = \frac{C_{Li}(t_j)}{\sum_i C_{Li,i}(t_i)\Delta t_i}$$

- And finally, the mean residence time was computed as :

$$t_s = \frac{\sum_i t_i C_{Li,i}(t_i)\Delta t_i}{\sum_i C_{Li,i}(t_i)\Delta t_i}$$

Normalized RTD curves of the reactor before and after washing of the interstitial biomass





Originally, volume of the reactor = 948 liters

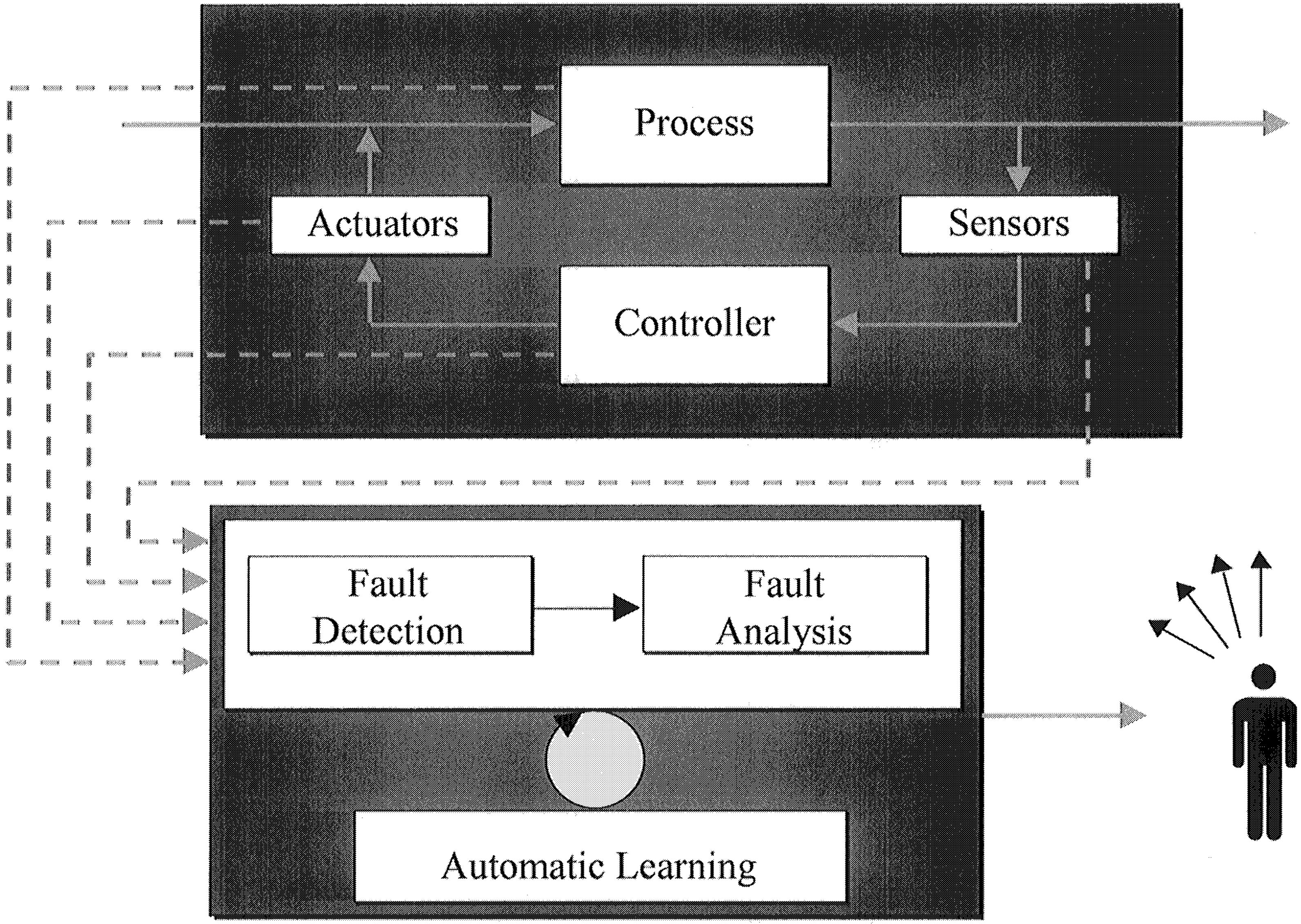
Before washing : 730 liters   After washing : 830 liters

Computed analytically : 595 liters

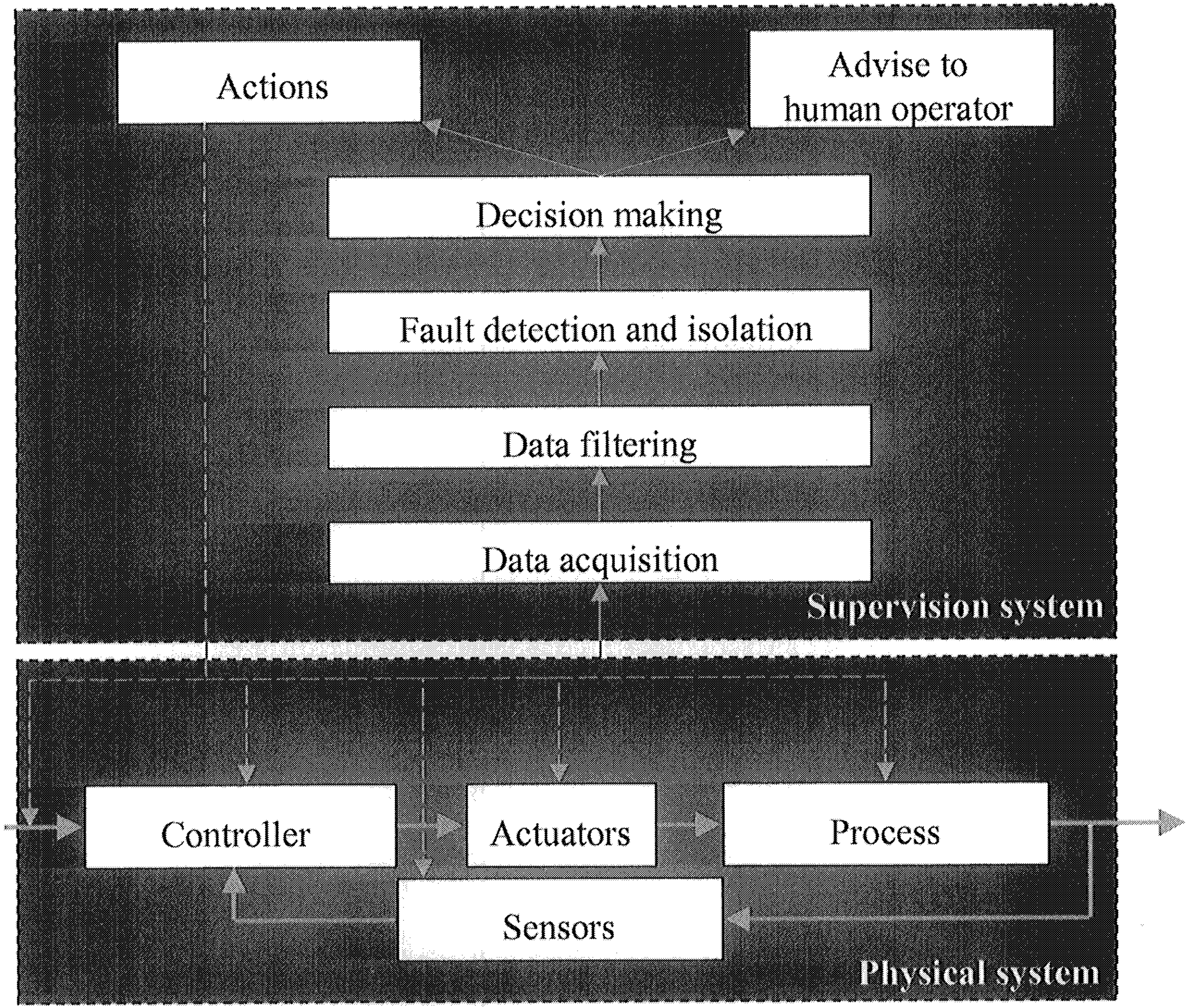
On-line Applications

- Principles
- FDI using simple analytical models
- FDI using heuristic knowledge
- FDI using data-based approaches

Integration of a diagnosis module in the control scheme

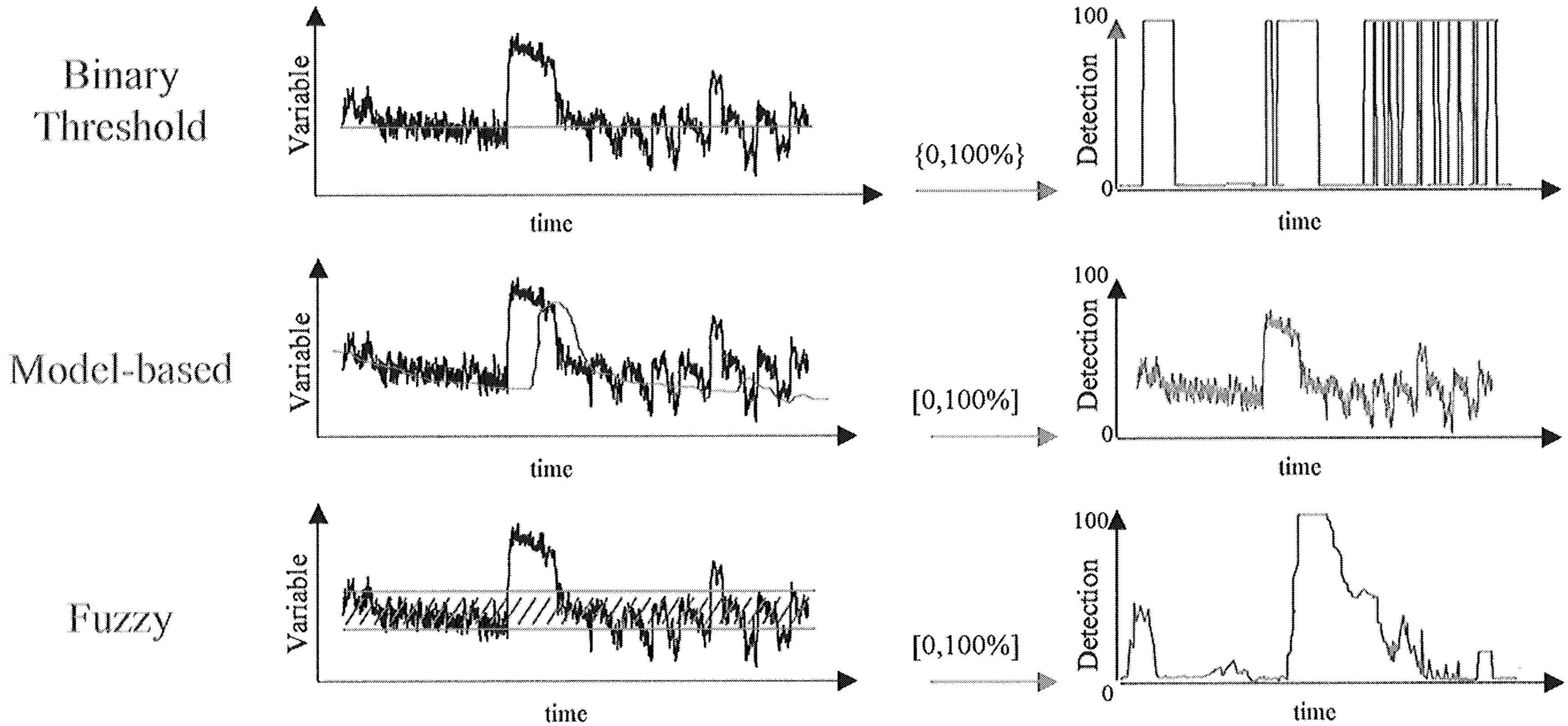


The different sub-tasks of a supervision system



Fault detection systems

Objectives
- High sensitivity (fast detection)
- Low false alarm rate
- Indication of the degree of urgency

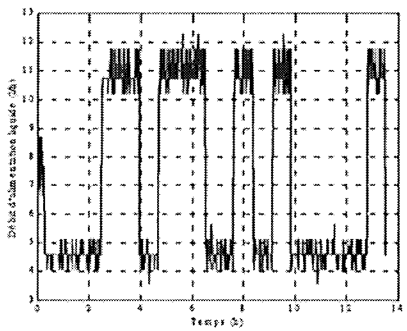


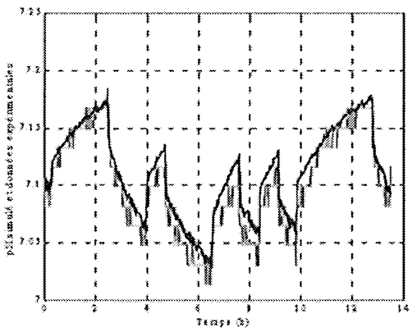
On-Line Diagnosis using Simple Analytical Models

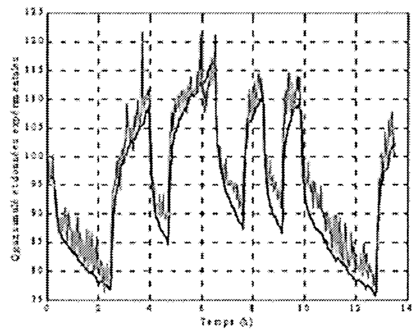
Using a 5^{th} order black box linear (ARMAX) model such as :

$$x(k+1)=\begin{bmatrix} 1{,}8081 & 1 & 0 & 0 & 0 \\ -0{,}8094 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0{,}3679 & 0 & 0 \\ 0 & 0 & 0 & 1{,}2268 & 1 \\ 0 & 0 & 0 & -0{,}2330 & 0 \end{bmatrix} x(k)+\begin{bmatrix} 0{,}5356 \\ -0{,}5241 \\ 0 \\ -0{,}005 \\ 0{,}0048 \end{bmatrix} u(k)$$

$$\begin{bmatrix} Q_{gaz}(k) \\ pH(k) \end{bmatrix}=\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} x(k)$$

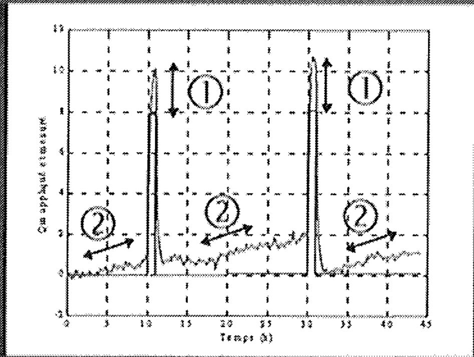


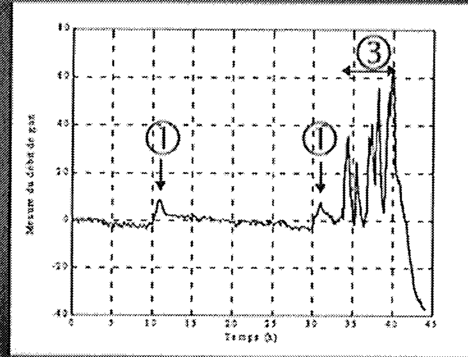




On-Line Diagnosis using Simple Analytical Models

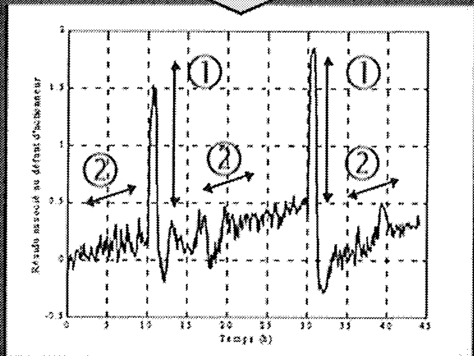
① and ② are defaults on the input liquid feed flow and ③ is a default on the biogas output flow



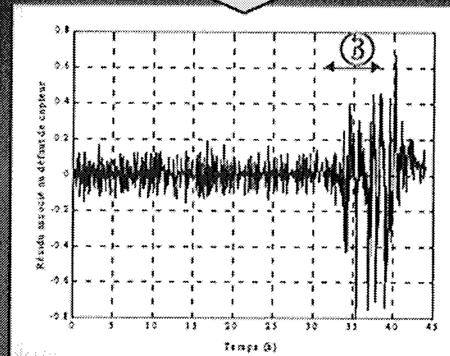


Measurements

Fault Residuals



Residual signal associated to the input liquid feed flow



Residual signal associated to the output biogas flow

Fault are perfectly detected and isolated !

Contents of the Presentation

INRA

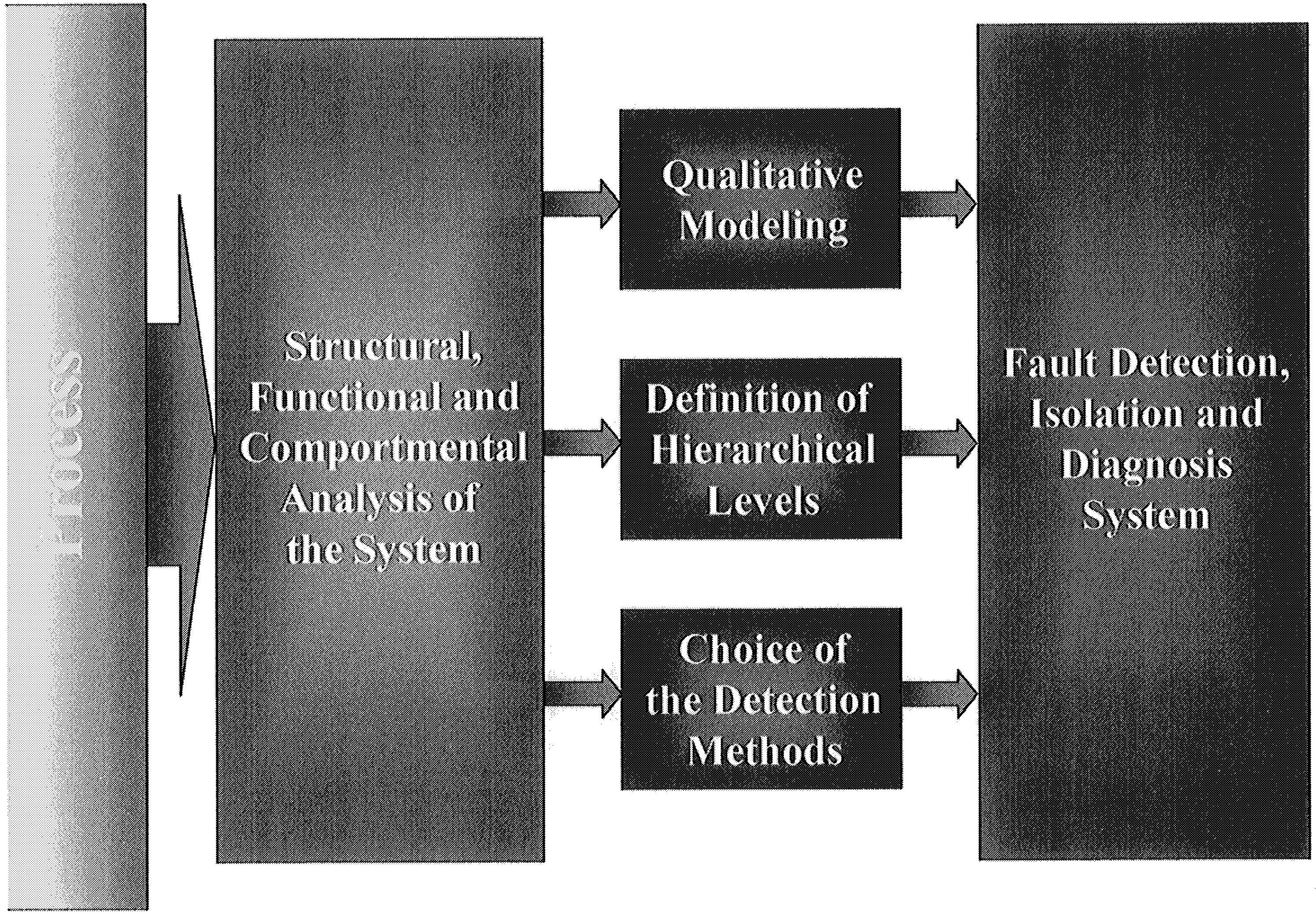
Hypothesis

We do not have any model...
We do not have large amount of on-line data ...

But we know how the process works (i.e., some expertise is available !)

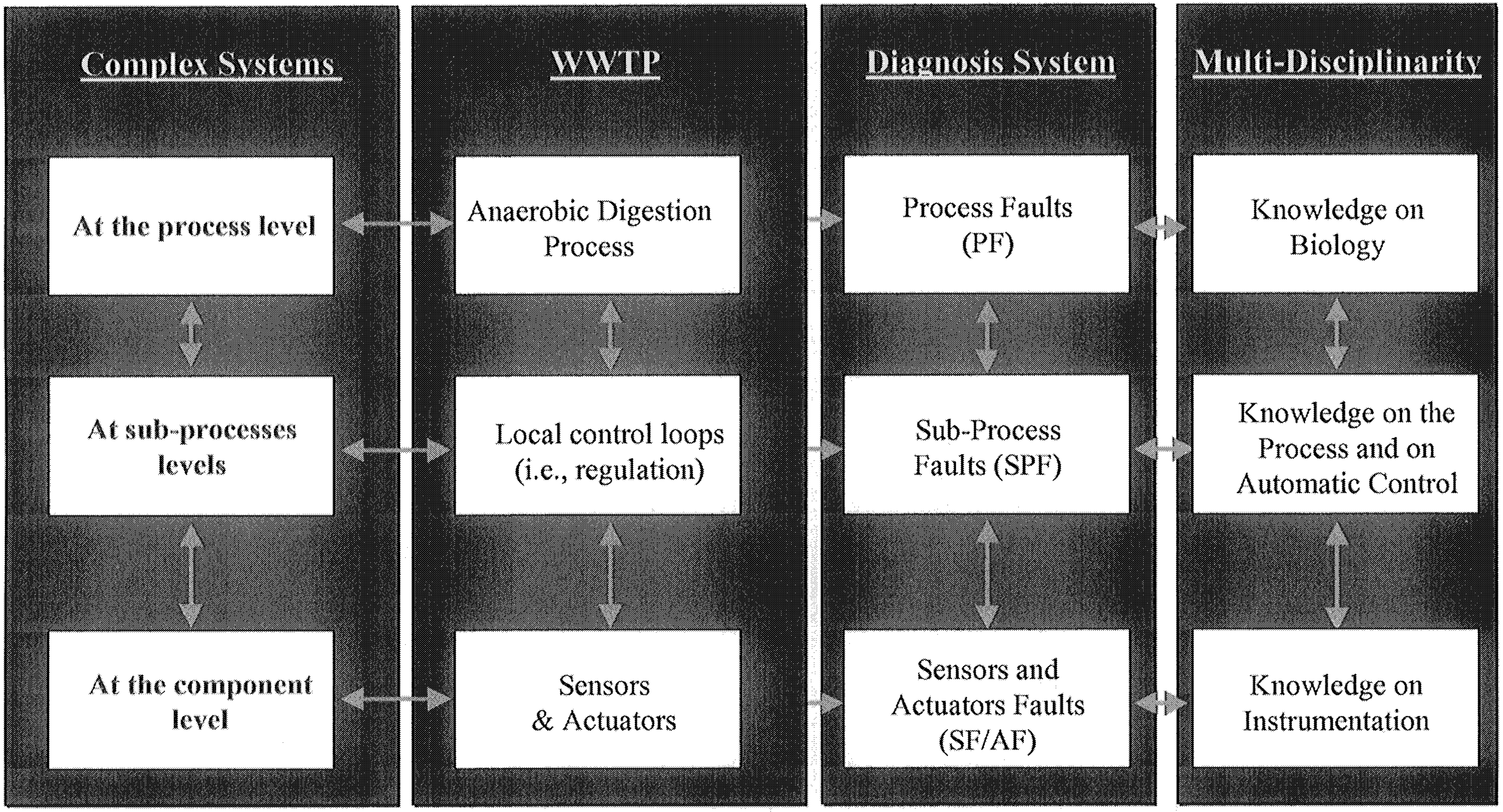
↳ ***Fuzzy Logic could be a solution***

Diagnosis based on Process Knowledge



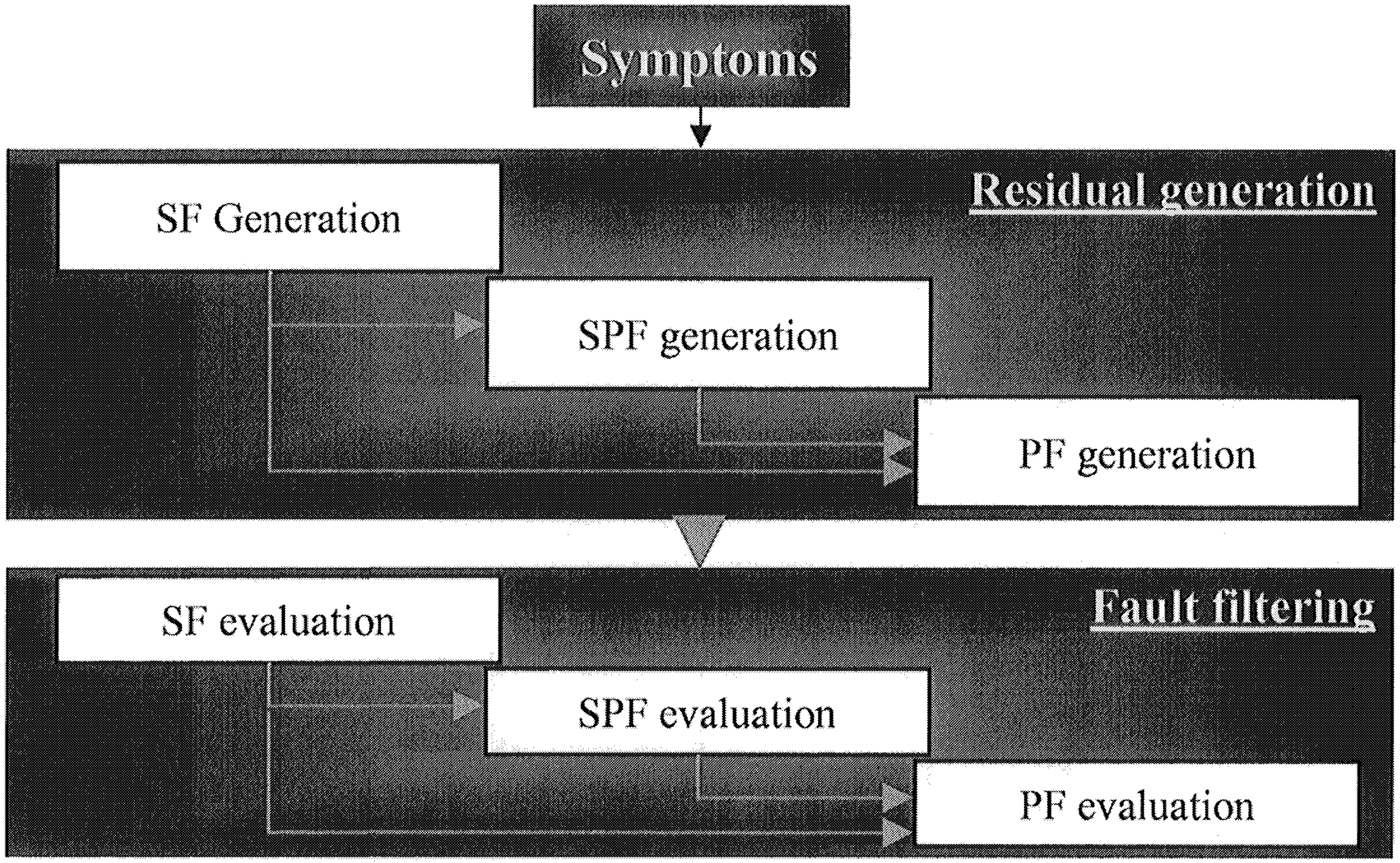
Heuristic based Diagnosis

↳ *Necessity of a fine (i.e., structure, function, behavior) and multi-disciplinary analysis of the process*

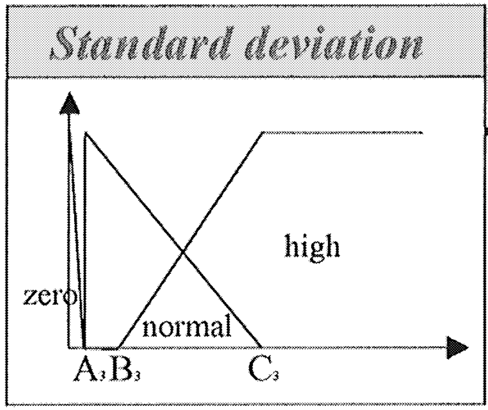


The Fault Detection and Isolation scheme

- **Sensor Fault (SF) :** Degree of confidence about the information given by the sensor
- **Sub Process Fault (SPF) :** Shows the presence of problems in local loops
- **Process Fault (PF) :** Indication of the importance of the SF and SPF faults on the overall anaerobic digestion process

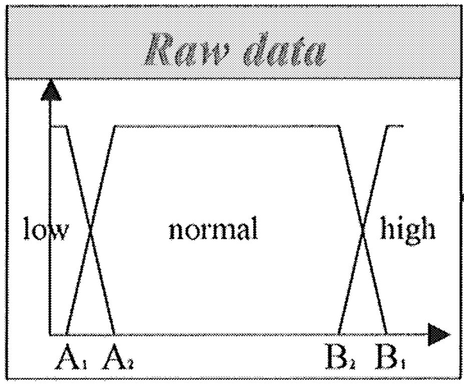


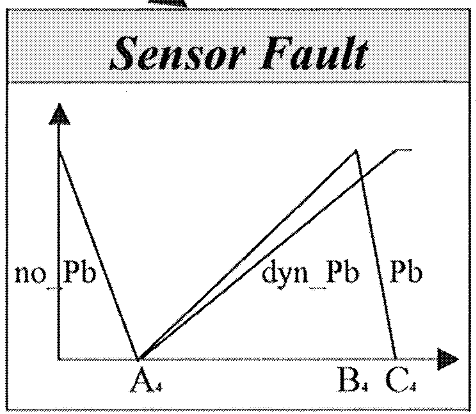
An Example of SF Fuzzy Residual Generation



Rules Base

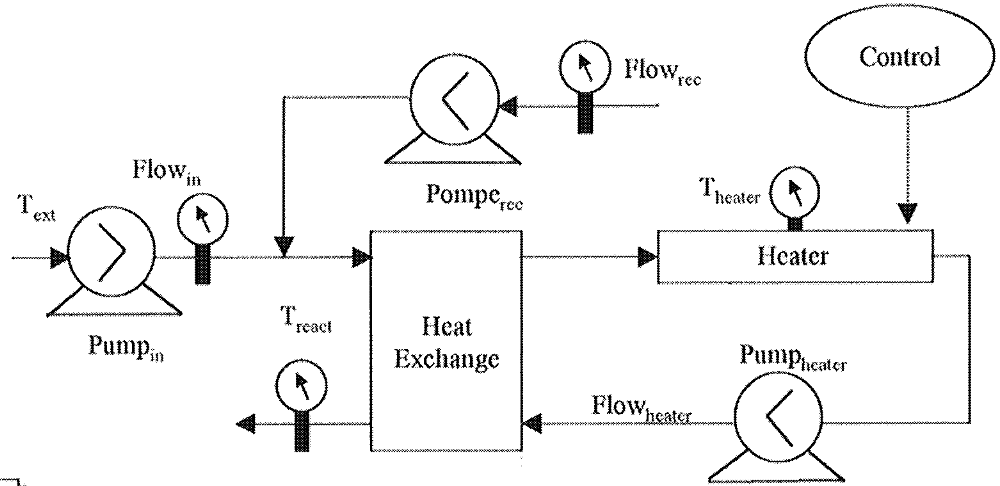
		Standard deviation		
		zero	normal	high
Raw data	low	*Pb*	*Pb*	*Pb*
	normal	*Pb*	*no_Pb*	*dyn_Pb*
	high	*Pb*	*Pb*	*Pb*





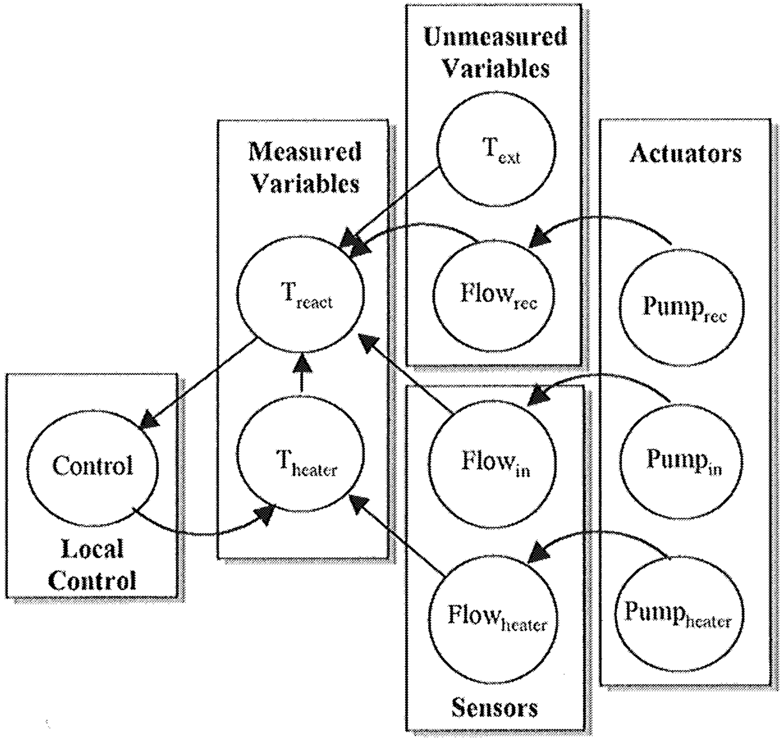
INRA Structural Analysis at Sub-Process and Process Levels

Formalisation of subprocess interactions

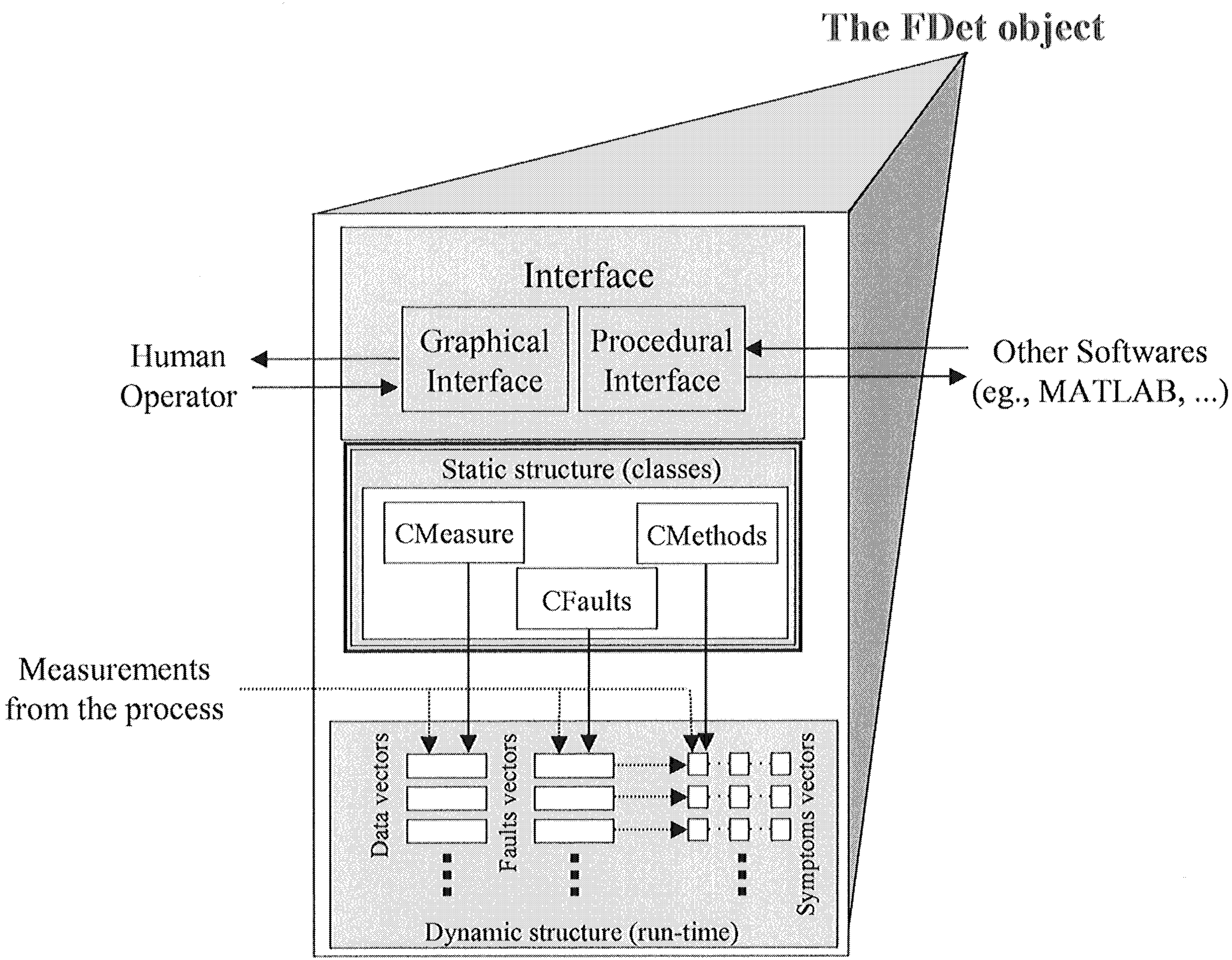


Temperature Sub-System

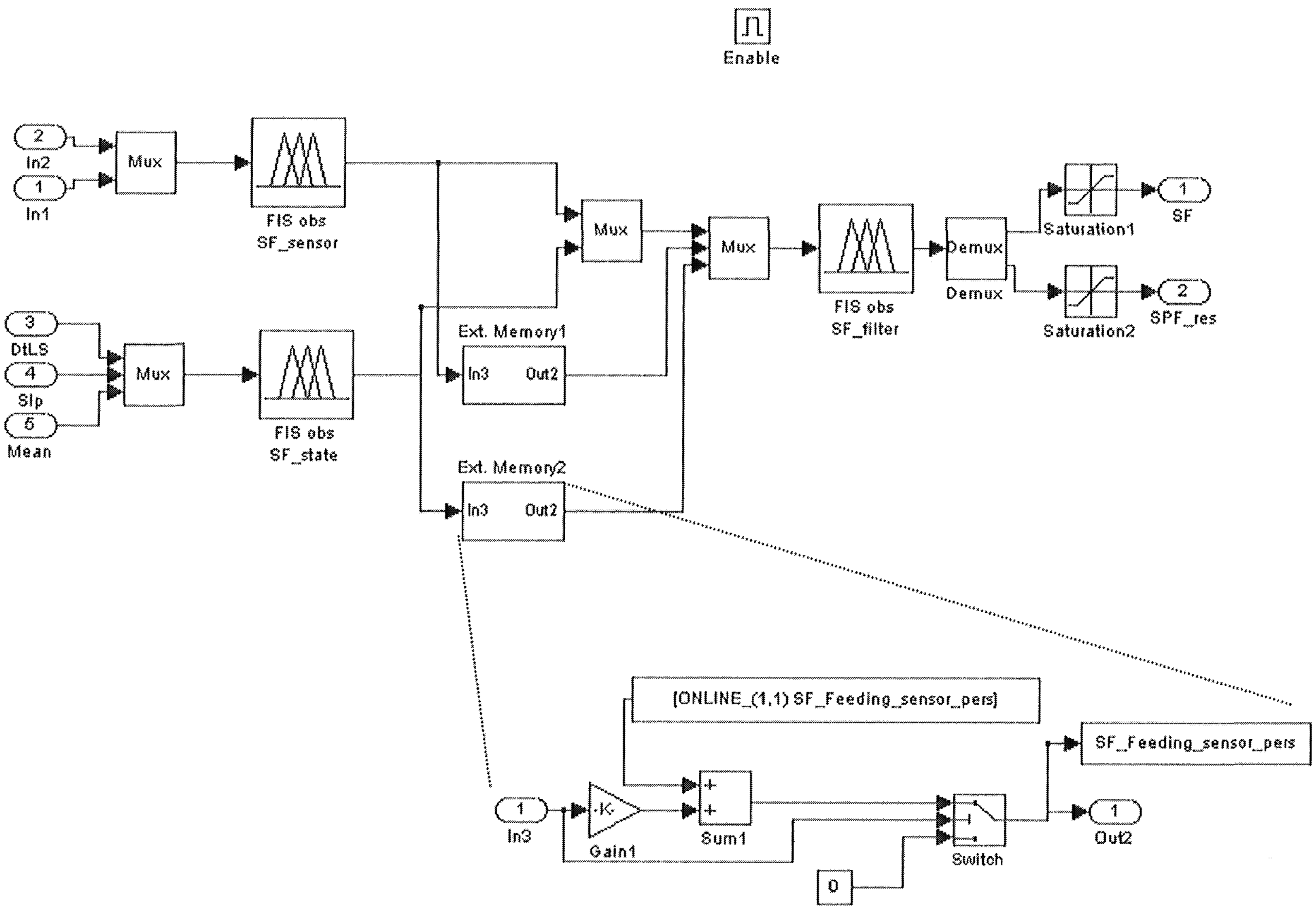
Graph of causal influences



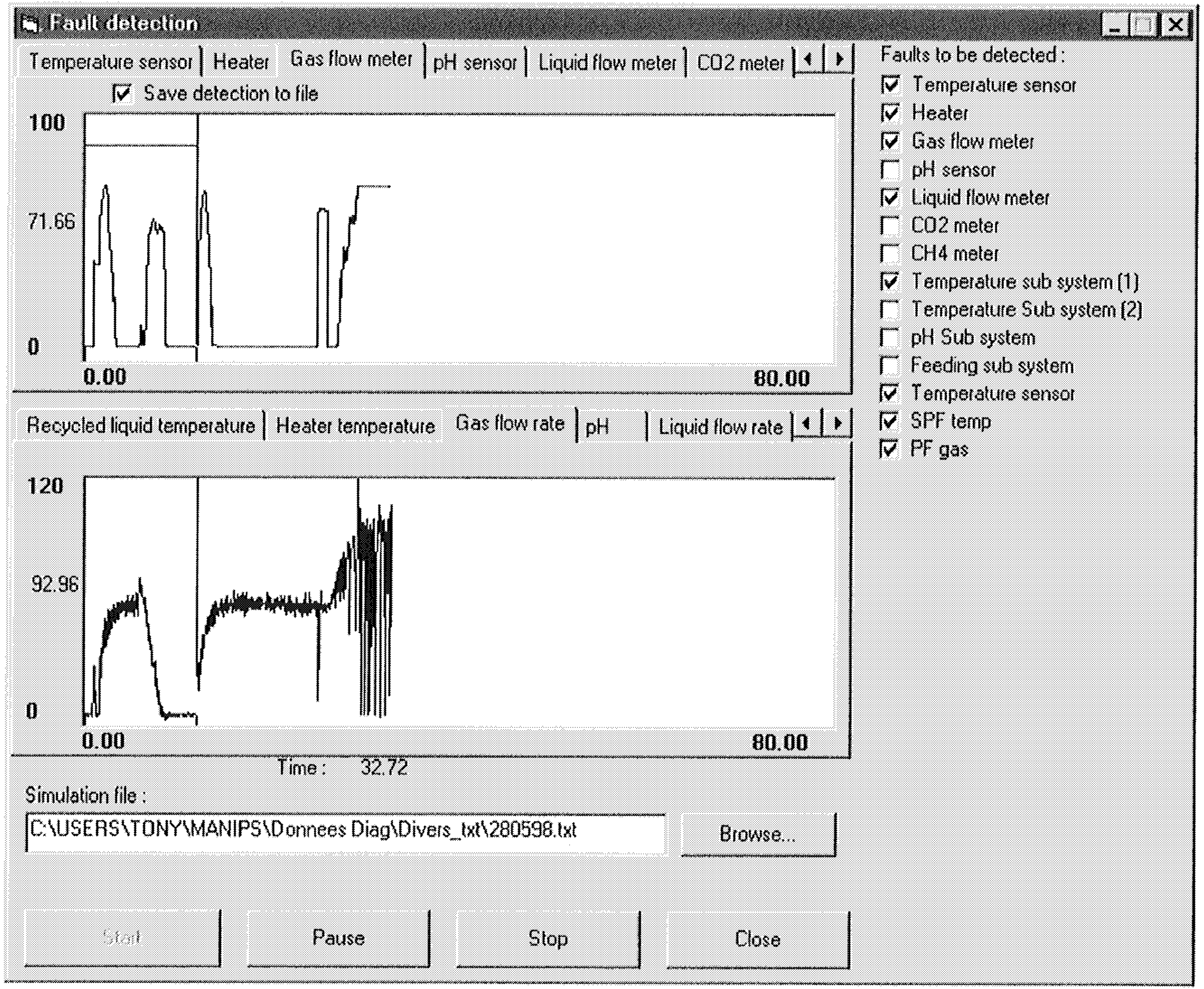
Internal structure of the fault detection module



Fuzzy module for SF detection and SPF residual generation



Graphical User Interface of the FDet Object

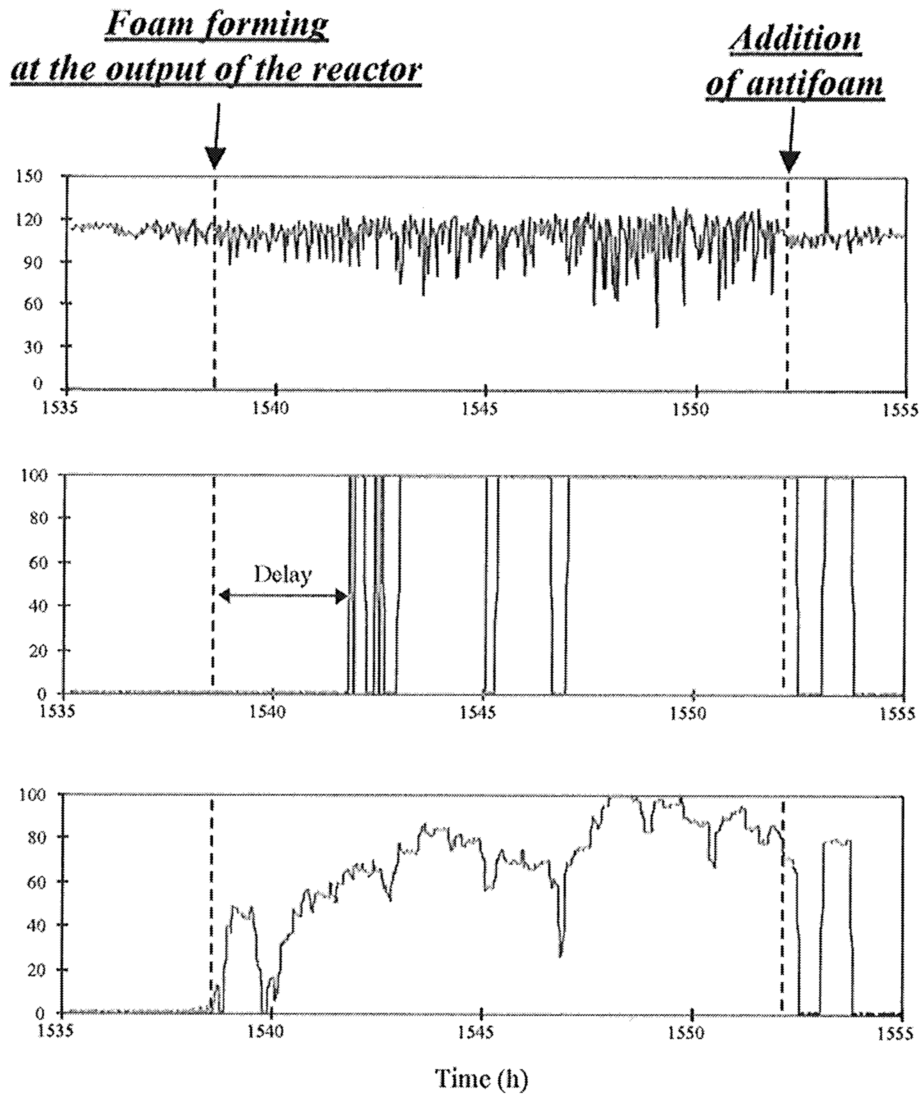


Comparison between Binary and Fuzzy Fault Detection

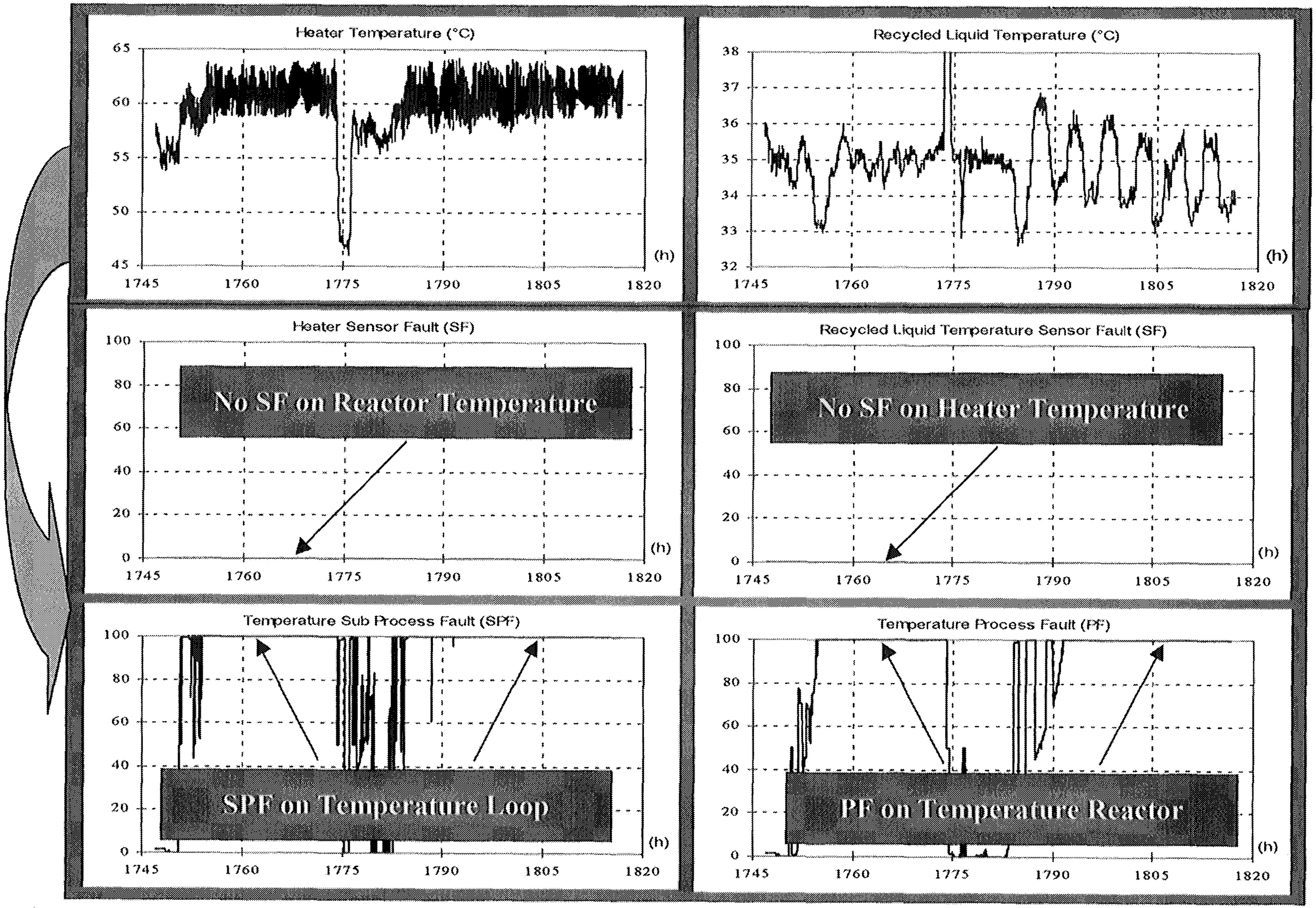
Output Gas Flow Rate (l/h)

Binary Fault Detection

Fuzzy Fault Detection

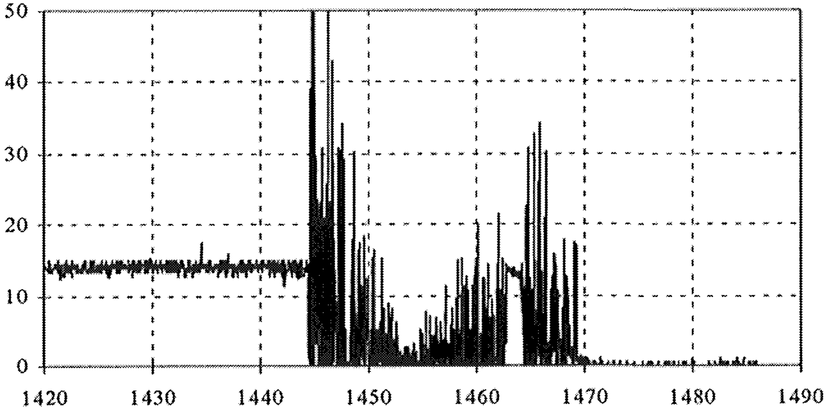


Fuzzy SF, SPF and PF for the temperature loop

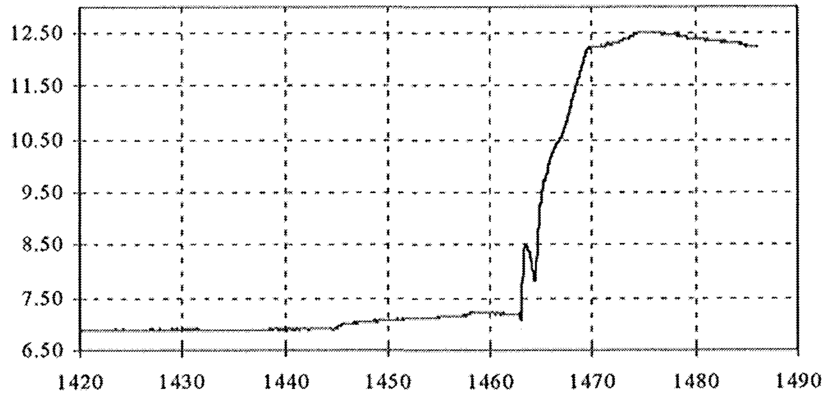


Fuzzy SF, SPF and PF for the feeding loop

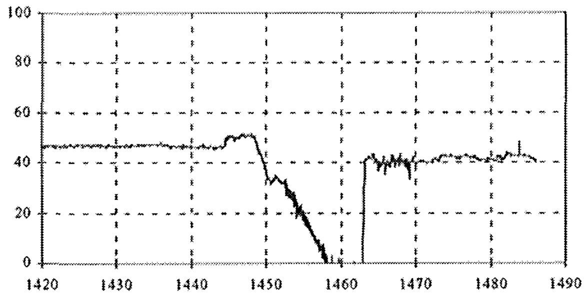
Input Liquid Flow Rate (l/h)



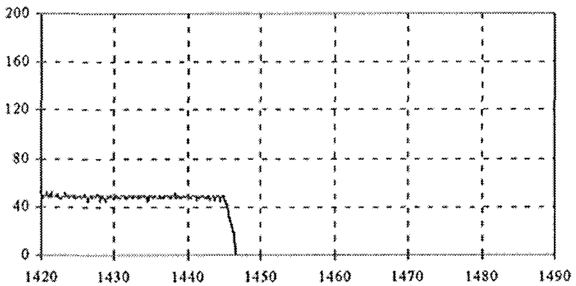
pH (UpH)



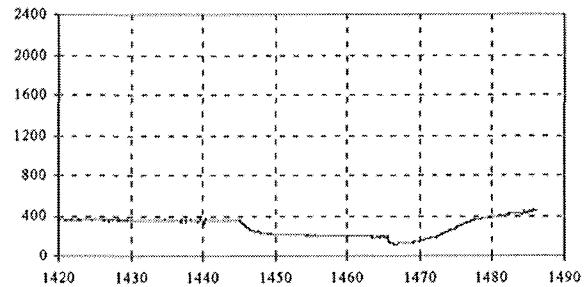
Recirculation flow rate (l/h)



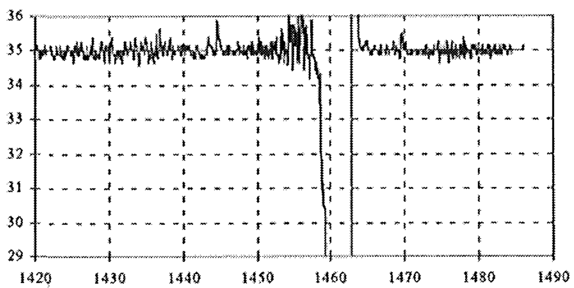
Output Gas Flow Rate (l/h)



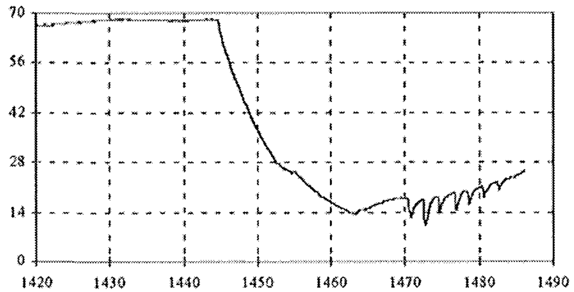
COT (ppm)



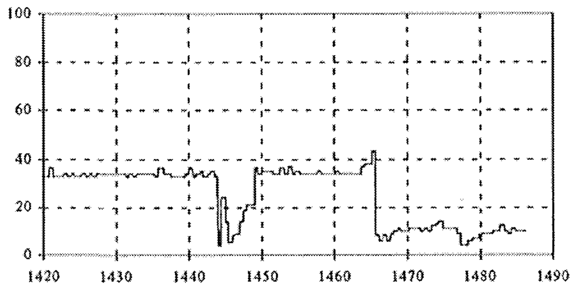
Recycled Liquid Temperature (°C)



CH4 (%)

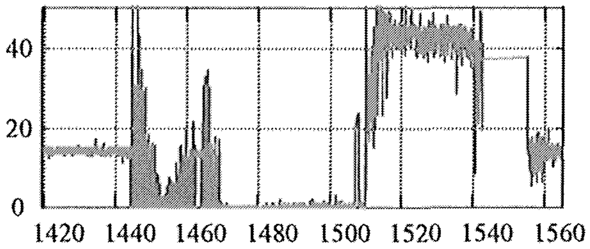


AT (méq/l)

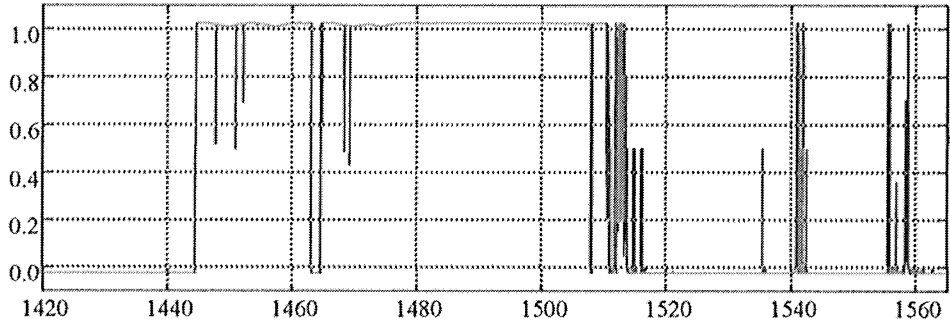


Fuzzy SF, SPF and PF for the feeding loop

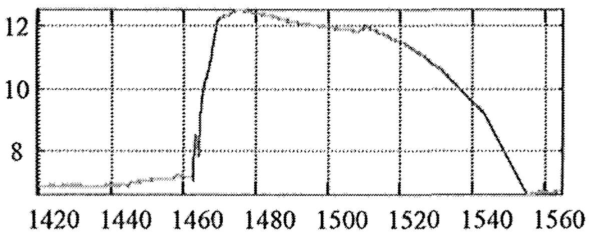
Qin signal



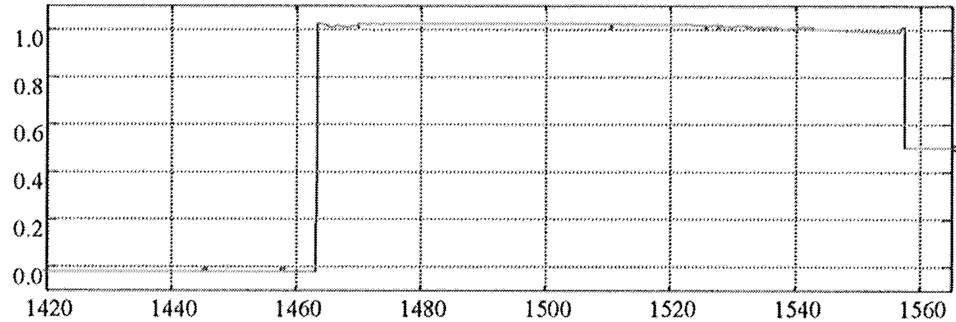
Residual from the SF/AF module related to Qin



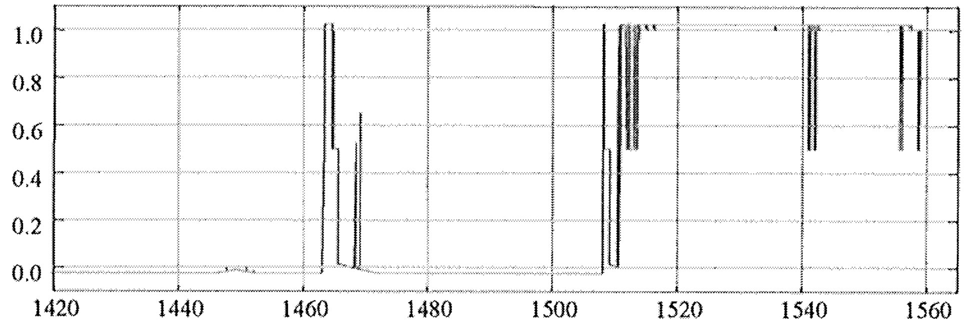
pH signal



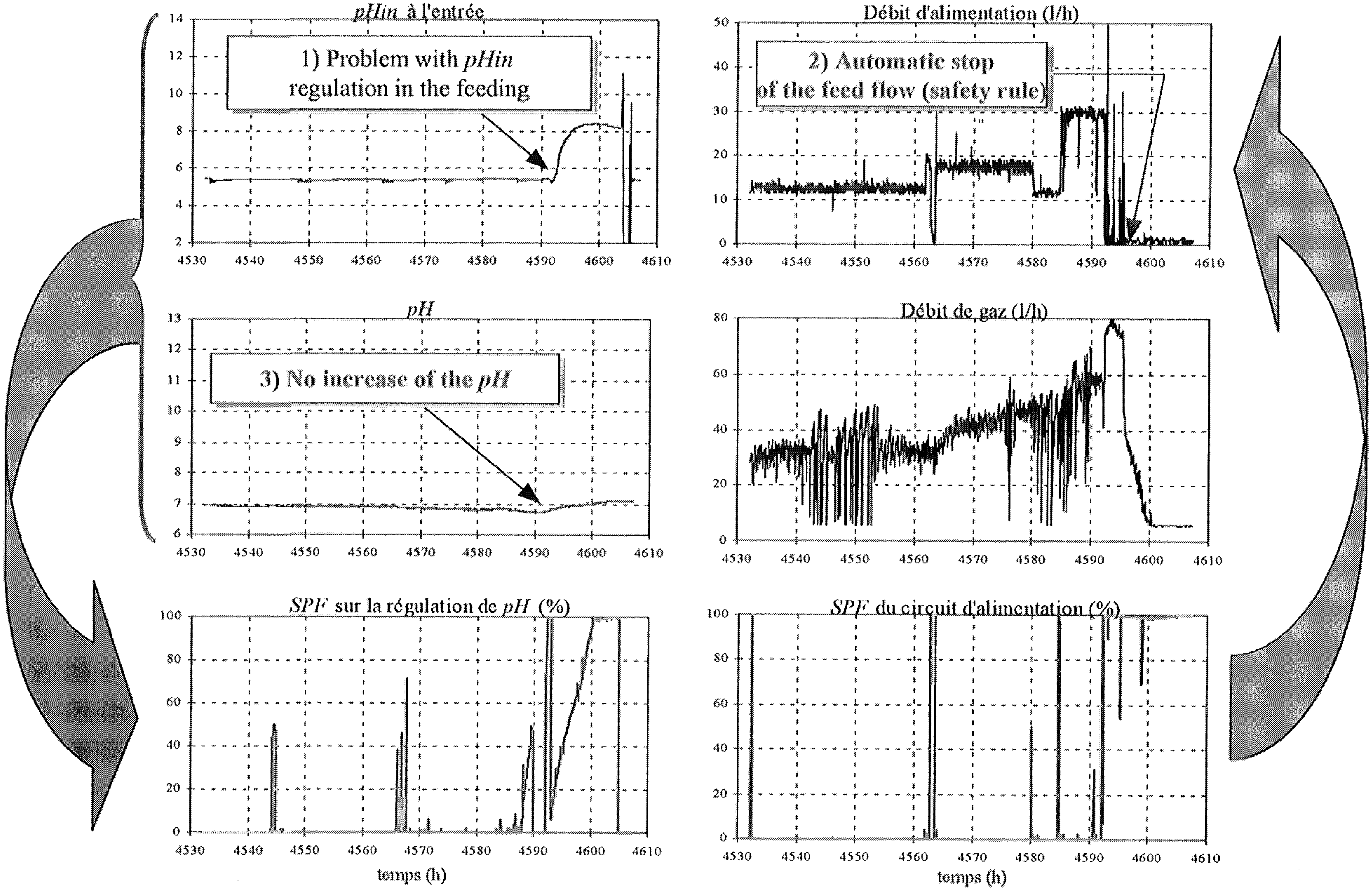
Residual from the SF/AF module related to pH



SPF related to pH (i.e., the two above residuals are aggregated)



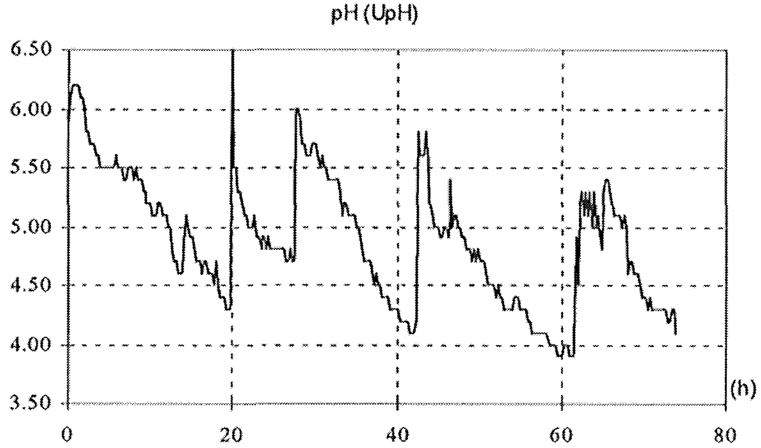
Closed Loop FDI

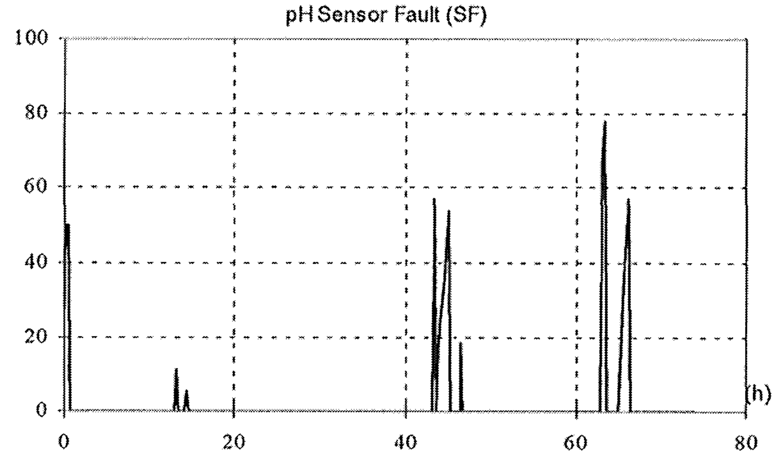


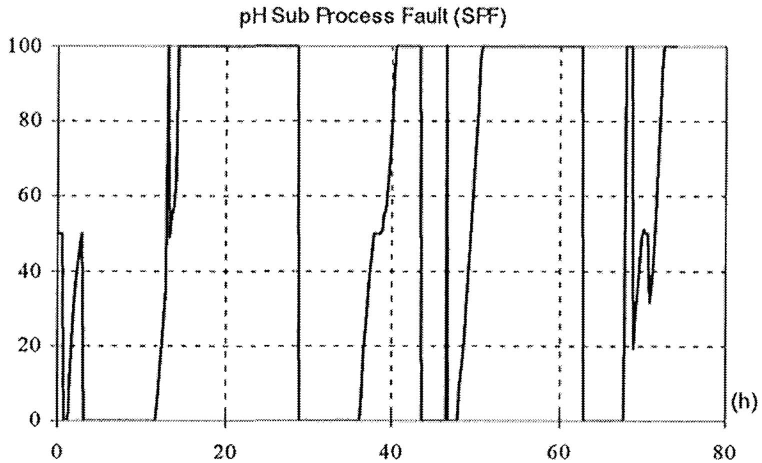
Time required before being back to normal : <u>24 hours</u> ! (vs. 1 month)

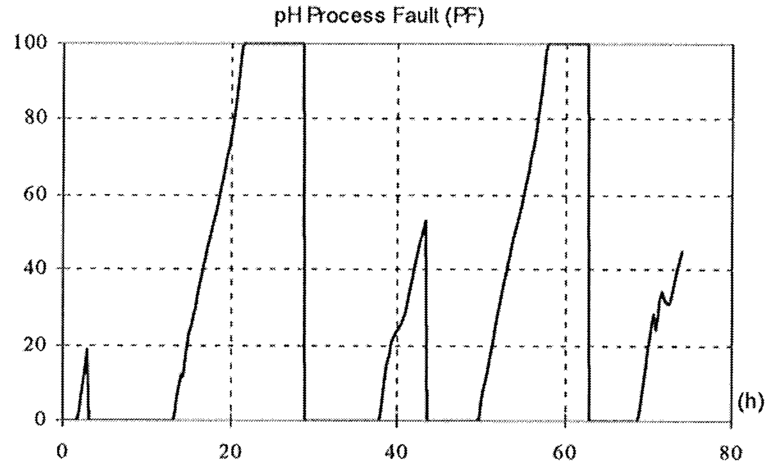
SF, SPF and PF on the pH "regulation loop"

At the industrial scale (other process, other wastewater)

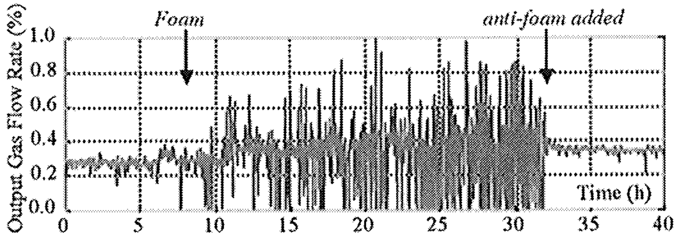








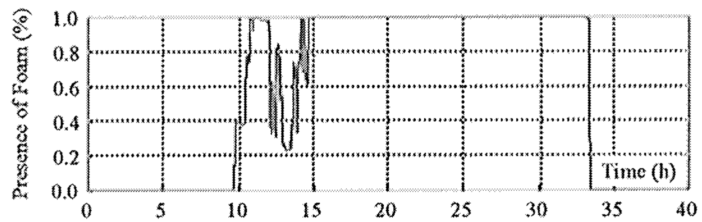
Generalisation of the approach



Measurement

Comparison to fuzzy thresholds

Degree of Fault



However, a wastewater treatment plant is anything but in stable conditions !

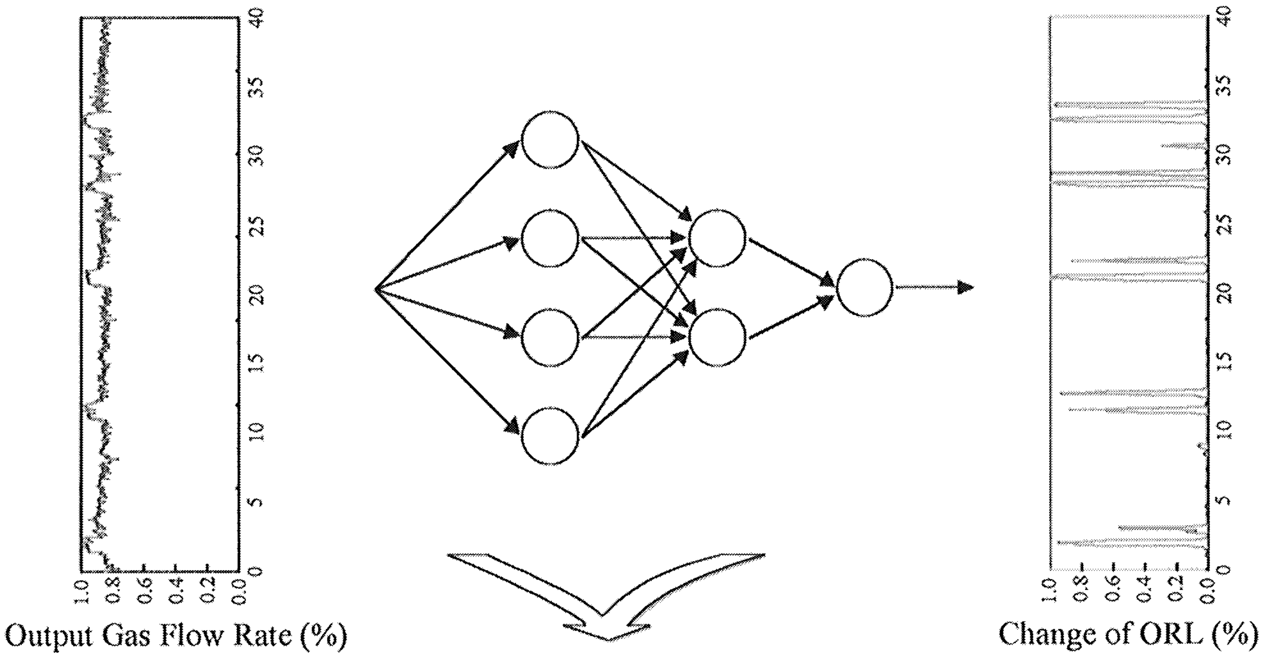
⇒ *It is thus necessary to adapt the thresholds to the practical situations*

We could have used again the Fuzzy Logic approach,

Instead, we used the Artificial Neural Networks.

In our Fault Detection System

- Each variable is connected to a specific Artificial Neural Network
- The training phase is performed based on the results obtained during the fuzzy fault detection previously achieved off-line.

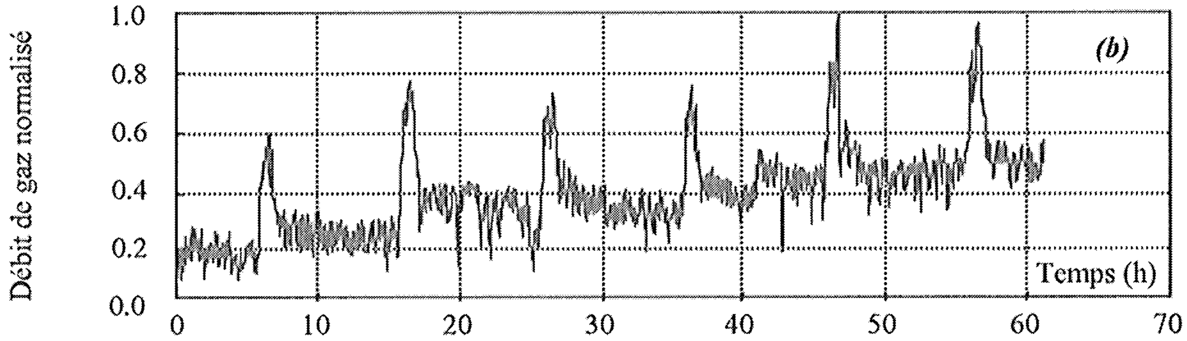


Tuning of the Artificial Neural Networks Parameters

In our Fault Detection System

- Once the training parameters have been determined, the Artificial Neural Network is used on-line as a non linear black box model between the measurements and the defaults.

- In other words, from measurements such as :



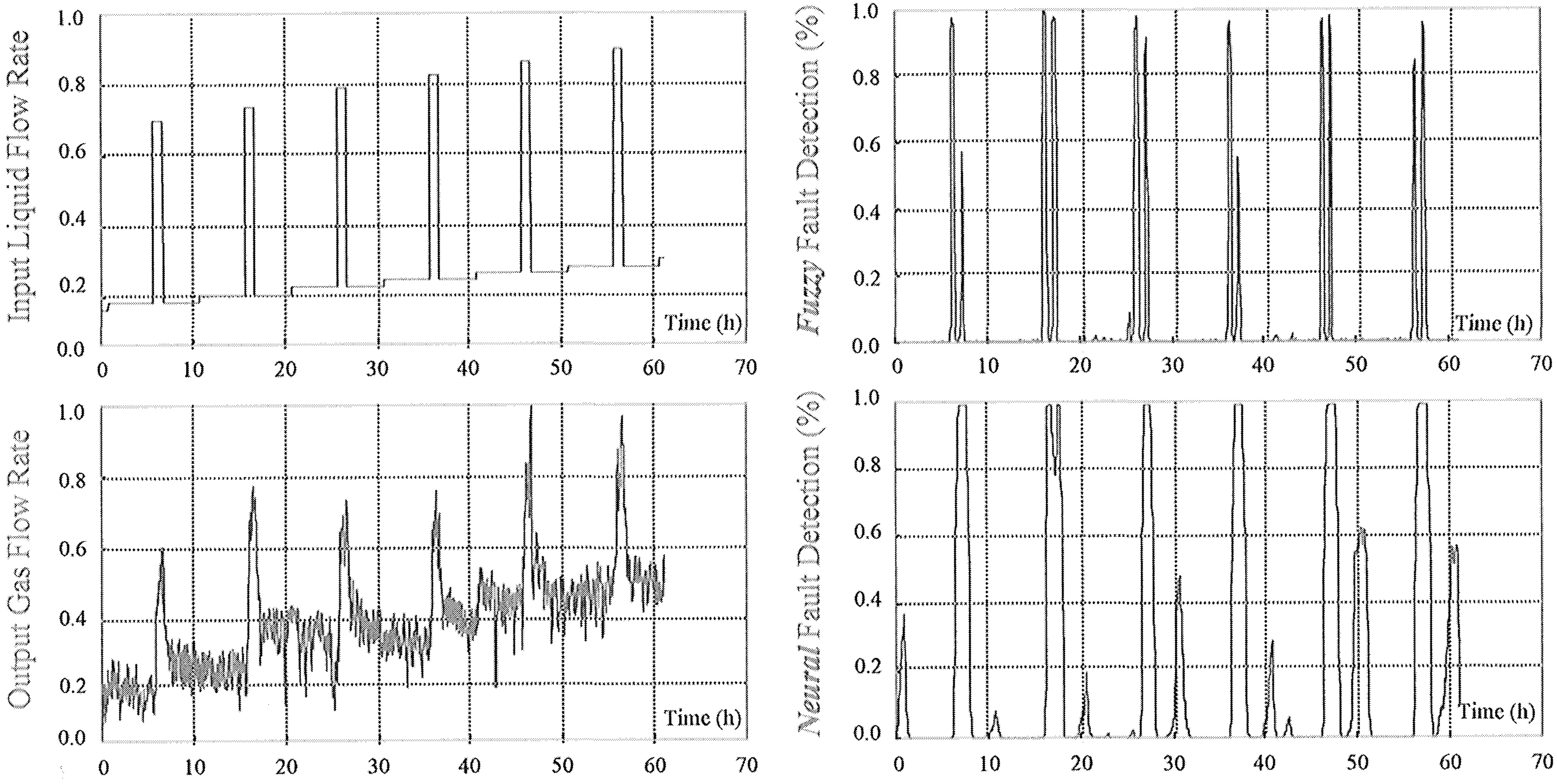
and that were not used during the training phase, the Artificial Neural Network indicates at each sampling time a new fault indicator.

Main avantage : Comparison to thresholds has disappeared !

Comparison between Fuzzy and Neural Fault Detection

Objective : Detection of Changes in Organic Loading Rate

(without measuring the input liquid flow rate)



INRA

Evaluation of the different FDI approaches

1) The **Fuzzy Logic** :

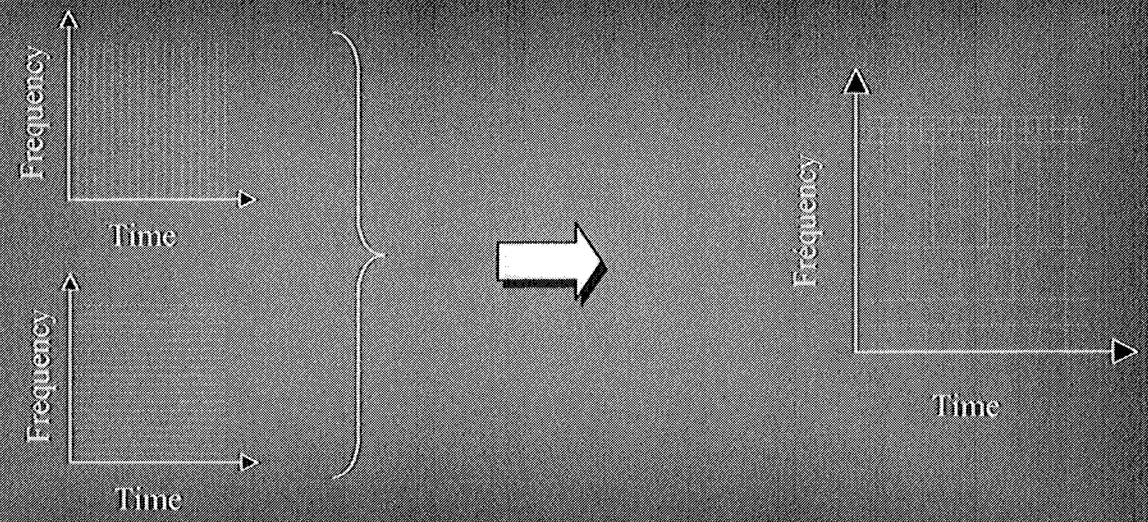
- Is ***"robust"*** for fault detection (i.e., the number of false alarm is minimized),
- Is ***very simple to use.***

2) The **Artificial Neuron Networks** :

- Are very well suited for ***on-line use,***
- Allow an ***automatic learning*** (through recursivity).

Data Based Diagnosis

Signal Processing



Classical Approaches :
analysis in time *or* in frequency

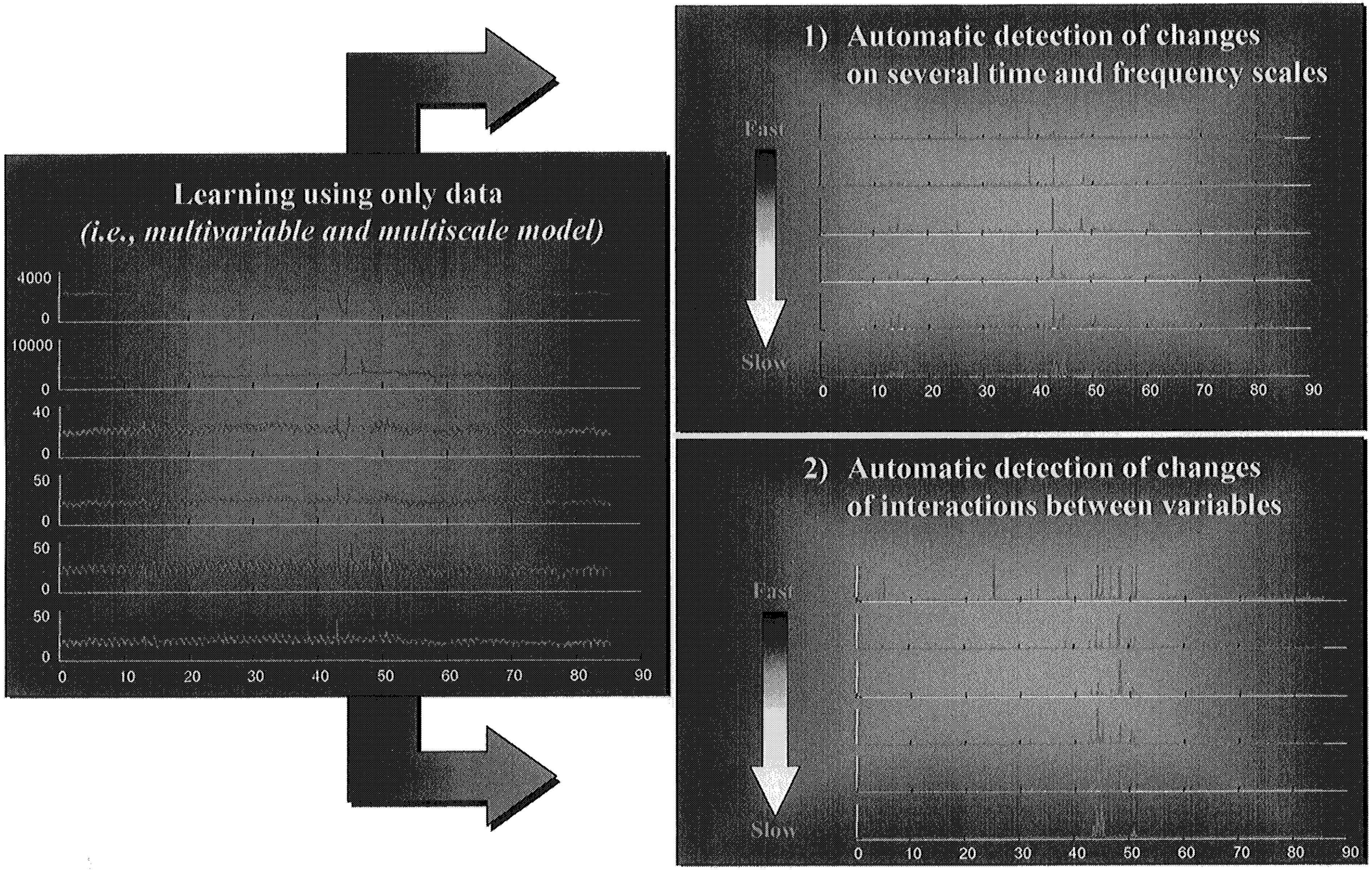
Using wavelets :
compact and multiscale analysis

Signal Analysis

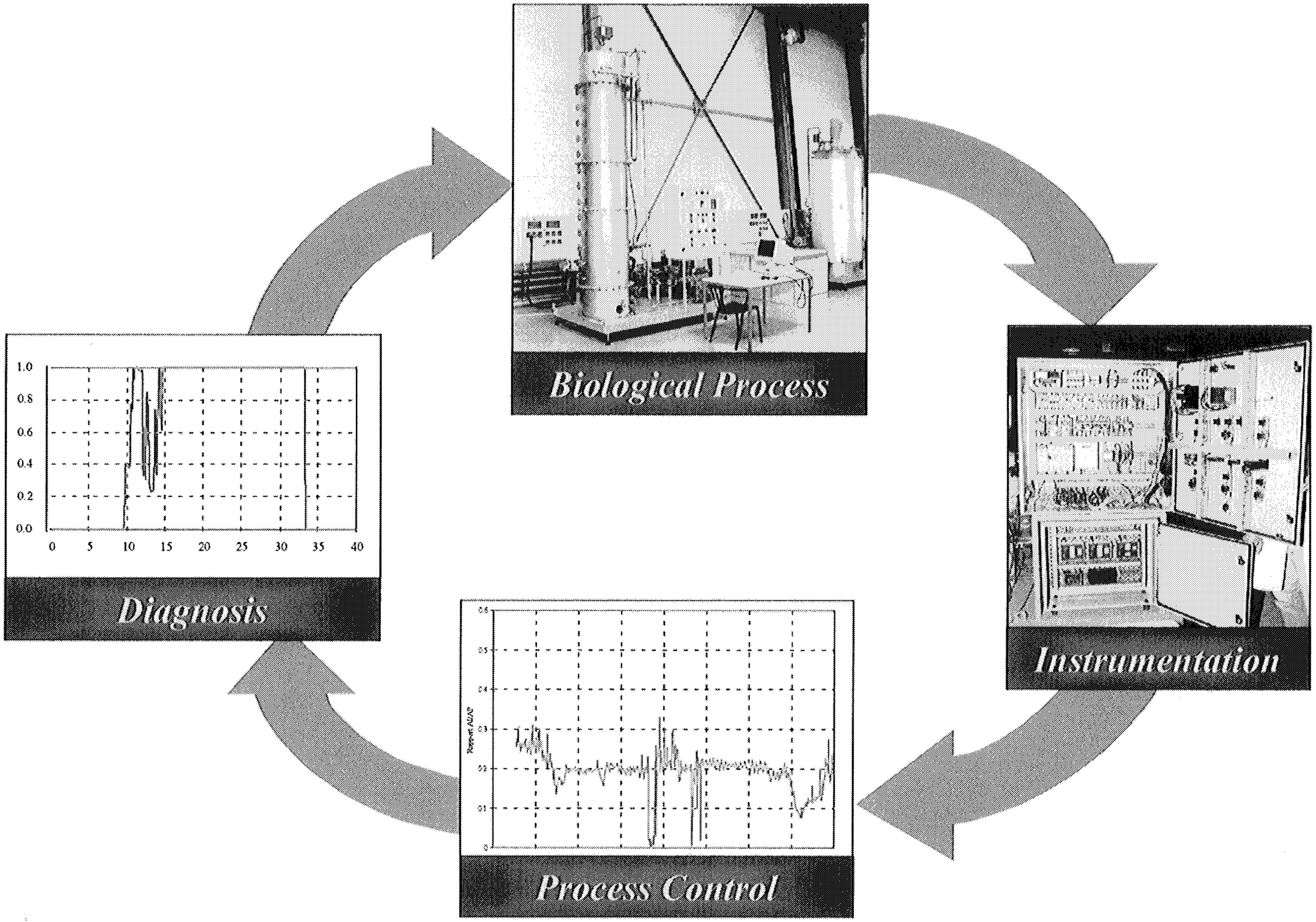
✓ Choice of variables by genetic algorithms

✓ Statistical linear (PCA, PLS) and non linear (ANN) classification

Data Based Diagnosis



Conclusions




THAT'S ALL FOLKS ...
THANK YOU VERY MUCH FOR YOUR ATTENTION !!!