HAL
open science

Acoustic monitoring of sodium boiling in a Liquid Metal Fast Breeder Reactor from autoregressive models

Issa Cherif Geraldo, Tanmoy Bose, Komi Midzodzi Pekpe, Jean-Philippe Cassar, Amiya Rajan Mohanty, Kévin Paumel

To cite this version:

Issa Cherif Geraldo, Tanmoy Bose, Komi Midzodzi Pekpe, Jean-Philippe Cassar, Amiya Rajan Mohanty, et al.. Acoustic monitoring of sodium boiling in a Liquid Metal Fast Breeder Reactor from autoregressive models. Nuclear Engineering and Design, 2014, 278, pp.573-585. 10.1016/j.nucengdes.2014.07.026 . hal-01298037

HAL Id: hal-01298037

https://hal.science/hal-01298037v1

Submitted on 5 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Acoustic monitoring of sodium boiling in a Liquid Metal Fast Breeder Reactor from autoregressive models

Issa Cherif Geraldo[a], Tanmoy Bose[b], Komi Midzodzi Pekpe*[a], Jean-Philippe Cassar[a], A. R. Mohanty[b], and Kévin Paumel[c]

[a] *Laboratoire d'Automatique, Génie Informatique et Signal (LAGIS UMR CNRS 8219), Université Lille 1, Sciences et technologies, Avenue Paul Langevin, BP 48, 59651 Villeneuve d'Ascq CEDEX, France.*
[b] *Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India.*
[c] *CEA, DEN, Nuclear Technology Department, F-13108 Saint-Paul-lez-Durance, France.*

Abstract

This paper deals with acoustic monitoring of sodium boiling in a Liquid Metal Fast Breeder Reactor (LMFBR) based on Auto Regressive (AR) models which have low computational complexities. Some authors have used AR models for sodium boiling or sodium-water reaction detection. These works are based on the characterization of the difference between fault free condition and current functioning of the system. However, even in absence of faults, it is possible to observe a change in the AR models due to the change of operating mode of the LMFBR. This sets up the delicate problem of how to distinguish a change in operating mode in absence of faults and a change due to presence of faults.

In this paper we propose a new approach for boiling detection based on the estimation of AR models on sliding windows. Afterwards, classification of the models into boiling or non-boiling models is made by comparing their coefficients by two statistical methods, multiple linear regression (LR) and Support Vectors Machines (SVM). The proposed approach takes into account operating mode informations in order to avoid false alarms.

Experimental data include non-boiling background noise data collected from Phenix power plant (France) and provided by the CEA (Commissariat à l'Energie Atomique et aux énergies alternatives, France) and boiling condition data generated in laboratory. High boiling detection rates as well as low false alarms rates obtained on these experimental data show that the proposed method is efficient for boiling detection. Most importantly, it shows that the boiling phenomenon introduces a disturbance into the AR models that can be clearly detected.

Keywords – Fault detection, Sound analysis, Nuclear plant, Autoregressive models, Classification methods.

*Corresponding author: Tel.: + 33 3 20 43 45 43 – Fax: + 33 3 20 33 71 89
Email address: `midzodzi.pekpe@univ-lille1.fr` (K. M. Pekpe)

Nomenclature

1_L	A vector of length L whose components all equal 1
a_i	A vector of consecutive values of the $i-$th ARLPC coefficient for a given sensor
a_w^r	The one-sensor AR model associated to the record by sensor r for signal w
a_w	The multi-sensor AR model associated to the signal w
$\bar{a}$	A mean multi-sensor model
$A_{i,\bullet}$	The $i-$th row of the A matrix
$[A \mid B]$	Concatenation of the A and B matrices
α	A real number in $[0, 1]$ representing a prevision interval risk
α_w	A real coefficient associated to signal w
B	A parameter matrix
d	The total number of environment parameters
e_w	A vector representing the non-explained part of a_w model
F	Frequency of the acoustic signals
L	Size of the learning set
N	Number of samples of a background signal
p	The total number of components of the multi-sensor models
p_r	The number of components of the AR models for sensor r
R	The total number of sensors
$s[n]$	The $n-$th sample of the signal s
$\hat{s}[n]$	The estimation of the $n-$th sample of the signal s
s_w^r	Acoustic signal of signal w recorded on sensor r
$(s_w^r)_{[k-p_r:k-1]}$	A vector of length p_r containing all the samples of signal s_w^r whose sample number lies between $k - p_r$ and $k - 1$
v, w	Signal numbers
x_w	A vector whose components are the values of the environment parameters of signal w
$*$	An optimal solution.

1 Introduction

Liquid Metal Fast Breeder Reactor (LMFBR) is highly efficient in power production. Liquid sodium can be used as coolant because of its high heat conductivity. As in every production plant, early fault detection method is needed for system safety. Indeed, the earlier a fault is detected, the smaller are the plant's damage. In a LMBFR, the reactor core is one of the main parts concerned with monitoring. The core of a fast reactor consists of fuel and reflector assemblies together with control rods. Fuel assemblies are placed in central section of the core and produce the fission reaction while control rods control the reaction rate. Heat produced in the fission reaction is carried outside by the coolant (that can be liquid sodium for example).

If for any unknown reason the heat removal from the fuel is degraded, the coolant temperature rises and coolant boiling might follow. This decreased heat removal can endanger the fuel integrity. Therefore, real time monitoring of the core temperature is necessary. Efficient methods based on temperature measurements are used for early boiling detection. In order not

only to strengthen but also to diversify the detection methods, acoustic method is proposed in this paper. This method is based on measurement of ambient sound record through wave-guide. The recorded data can further be analysed by different methods such as wavelet analysis, Power Spectral Density (PSD) and autoregressive modelling.

Carey (1979) had investigated the factors influencing background noise of a reactor. Data were collected over 17 months period and then analysed to determine the linear average spectra, auto and cross power spectra. Acoustic sound power is estimated and compared with results. It was shown that flow velocity influences acoustic power level and noise field can be considered as weakly stationary. Hayashi et al. (1996) developed a twice squaring method for real time sodium boiling detection. In their method, signal to noise ratio is enhanced by non-linear amplification of a band limited signal. Band-pass frequency is selected from PSD graphs, focussing on pulsive nature of boiling signal. It consists of five steps: band-pass filtering, squaring, another band-pass filtering and squaring and integration. A low pass filter is then applied to obtain the feature signal. The threshold for boiling detection is later calculated from the mean and the standard deviation of the feature signal in non-boiling conditions. This approach is successful if the mean and the standard deviation are always the same for signals. But if not, it sets up the problem of choosing an adequate value for the detection threshold.

Another method proposed by researchers is autoregressive model-based detection techniques. Hayashi (1997) has used Auto-Regressive (AR) models for sodium leak detection. He assumed that in normal functioning conditions, the background signal is stationary and then the prediction error from the AR model follows a Gaussian distribution. He has shown that prediction error deviates from gaussian distribution in non-normal functioning conditions. Inujima et al. (1982) also worked on boiling detection by analysing residual time series data of autoregressive model. In both approaches (Hayashi, 1997; Inujima et al., 1982), the selection of the reference AR model is difficult when the AR models obtained from different signals do not have the same characteristics. Indeed, the AR models change with the operating modes of the LMFBR. In order to avoid false alarms, it is then critical to distinguish a change in the operating mode in absence of faults and a change due to faults.

In this paper, we present a new approach for boiling detection based on classification of the AR models parameters estimation. We propose to divide background signals into multiple non-overlapping time windows and then estimate the AR models on each. The choice of AR model is motivated by the online implementation of the proposed method which imposes a low computational complexity method. There after models can be classified into boiling and non-boiling models by multiple linear regression and support vector machines.

The paper is organized in 5 sections. Section 2 presents the experimental data, deals with fault condition data generation and data preprocessing. In section 3 two boiling detection methods are proposed. Section 4 provides the results of the methods on the experimental data. Summary and conclusion are outlined in section 5.

2 Presentation of the experimental data

2.1 Nuclear Plant acoustic background noise recording in fault free-case

The original data furnished by the CEA consist of records of the acoustic background noise of the Phenix nuclear power plant made in 2009 (see Paumel et al. (2013) or Dall'Ava et al. (2009)). The data are recorded in normal functioning of the nuclear plant without any boiling. The data

are recorded by two wave guides called sensors 1 and 2 in the paper. The original sampling frequency is 500 kHz. The records were performed for different operating modes (or environment parameters). Seven environment parameters listed in table 1 were considered for the test.

Table 1: List of the environment parameters considered for the Phenix nuclear power plant

Parameter	**Unit**
Power of the reactor	MW
Inlet temperature of the core	°C
Outlet temperature of the core	°C
Primary pump 1 speed	RPM
Primary pumps 2 and 3 speed	RPM
Secondary pumps speed	RPM
Overpressure	mB

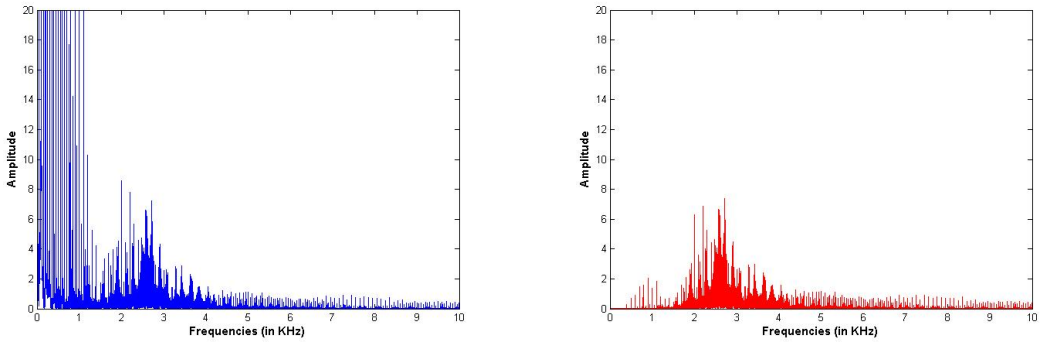
2.2 Data preprocessing

An example of raw signal spectrum for frequencies lower than 10 kHz is shown in figure 1(a). In no boiling condition, the background signal is made of noise from different sources like liquid coolant cavitation, vortex flow, shaft vibration and mechanical pump noise. Each part of the background noise could disturb the boiling detection thus the signals generated by these sources should be filtered in order to improve boiling detection. The pumps noise are the most energical part of the background noise. Fortunately, as it can be seen in figure 1(a) the most energical contribution of pumps noises are under 1 kHz and the sodium boiling acoustic noise frequencies lie between 1 kHz and 100 kHz (Hayashi et al., 1996). Therefore, the frequencies under 1 kHz must be filtered. In order to do that a 5−th order Butterworth filter with 2 kHz cut-off frequency is applied on the raw signal for removal of the most energetical parts of the pump noise. This cut-off frequency was chosen to make sure to remove the frequencies below 1 kHz. The spectrum of the filtered signal is shown on figure 1(b). It can be seen that the high amplitudes in low frequencies upto 2 kHz have been removed. The filtered signals are used in the rest of the paper. We assume the non-existence of cavitation, vortex and shaft vibrations in the LMFBR acoustic background noise during the experiments.

2.3 Boiling data generation

Background noise data supplied by CEA contains only records during normal functioning (non-boiling) of the power plant. But to check the efficiency of our proposed method, background noise in boiling conditions are needed. As these data are not available, one could record liquid sodium boiling sound and inject it into the non-boiling background noise to obtain boiling conditions background noise. But it is very complicated to boil sodium (at a temperature near 900 °C) in steady and controlled conditions. This experience would require a huge investment for the re-appropriation of the boiling sodium techniques.

It could then be interesting to substitute liquid sodium for another liquid for which the boiling experience will be simpler to perform. Bomeluberg (1968) has investigated applicability of water model testing instead of liquid metal tests by dimensional analysis. He has established relationship between cavitation, vortex and other flow problems in liquid metal with that of water.



(a) Raw signal spectrum

(b) Filtered signal spectrum

Figure 1: Comparison of the spectrum of raw signal and filtered signal for frequencies lower than 10 kHz.

Water is suggested as potential substitute of liquid metal with sufficient degree of accuracy when hydraulic characteristics are concerned. Prakash et al. (2011) have also used water instead of sodium for hydraulic experimental studies on a fast breeder reactor due to similar hydraulic characteristics of sodium and ease of testing. But one must be careful as far as the heat transfer is concerned.

The researches developed by the CEA (see Vanderhaegen et al. (2013)) show that complete thermal hydraulic and acoustic scaling of a sodium boiling loop to a water boiling loop is impossible. Therefore, a perfect transposition of the boiling sound source is impossible due to the large difference between the thermal diffusivity of water and sodium. However a partial transposition with a scale reduction factor and a power reduction factor can be assumed possible. To generate the background noise in boiling conditions, we have recorded water boiling sound afterwards we have injected it into the background noise supplied by the CEA.

The water boiling experiment was carried out at the Indian Institute of Technology Kharagpur. It was done in a steel container of 15 cm diameter and 8.5 cm height placed above an electric heater of power capacity 1000 W. Experimental set-up is shown in Figures 2 and 3. The transducer used was a B&K 4136 1/4" condenser microphone with a range of 4 Hz to 70 kHz. The sensitivity was 1.6 mV/Pa. This was used with a B&K 2804 microphone power supply which provided a 200 V polarisation voltage to the B&K 2669 microphone pre-amplifier. The high frequency boiling noise was acquired using a B&K PULSE analyzer with sampling frequency of 262144 Hz. The overall ambient noise level in the laboratory before the boiling took place is 28 dBA.

The boiling sound was recorded at four different stages depending on the speed of boiling. The first two stages correspond respectively to the two stages of nucleate boiling i.e. isolated bubbles regime and slug and columns regime. The third and the fourth ones correspond respectively to transitional boiling and film boiling as described in (Lienhard, J. H. IV, Lienhard, J. H. V, 2000).

2.4 Boiling condition data generation

We assume that the acoustic effect of the boiling is additive. As the boiling record sampling frequency is lower than the background noise frequency, the sampling frequency of the background

Figure 2: Boiling Experiment Set-up

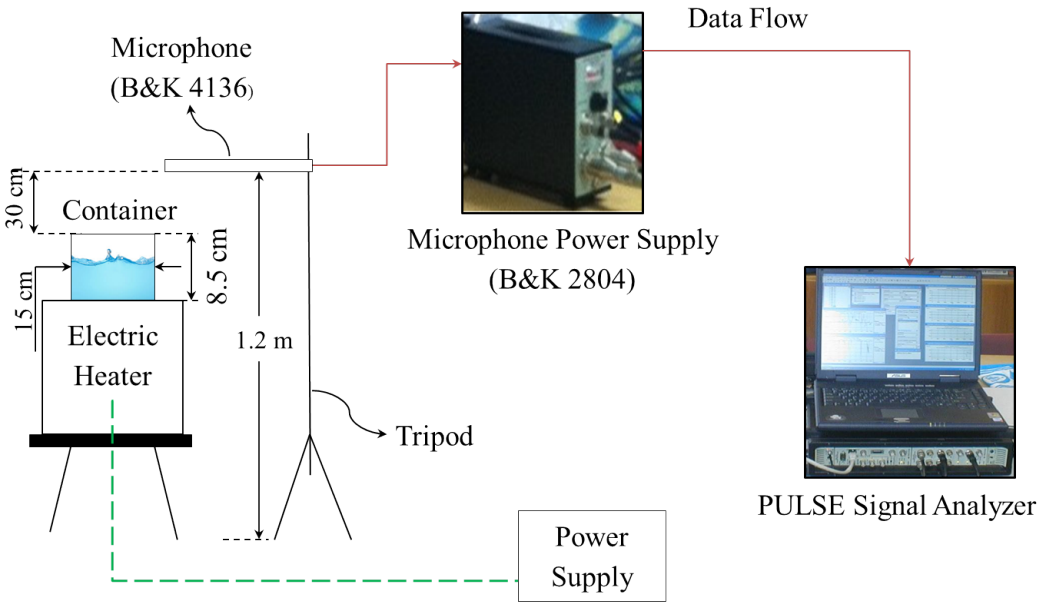


Figure 3: Line diagram of the experimental Set-up

signal has been reduced to 262 144 Hz.

Let s_s be an example of background noise of the nuclear power plant in non-boiling op-

eration and s_b be the boiling signal of common sampling frequency $F = 262\ 144$ Hz. Let $M = \min(N_s, N_b)$ where N_s is the number of points of s_s and N_b is the number of points of s_b. Let $\mathbf{s}_{sn}$ and $\mathbf{s}_{bn}$ the signals of duration $\frac{M}{F}$ seconds obtained respectively from s_s and s_b after taking the first M points and then normalizing by dividing each by its standard deviation. The output signal is generated by the following formula:

$$s_{\beta,t}[n] = \begin{cases} \mathbf{s}_{sn}[n] & \text{if} \quad n < m \\ \\ (1-\beta)\mathbf{s}_{sn}[n] + \beta \mathbf{s}_{bn}[n-m+1] & \text{otherwise} \end{cases} \tag{1}$$

where $m = \lfloor F \times t \rfloor$ (the integer part of $F \times t$) is the number of the sample corresponding to t, the time of the onset of boiling ($0 \leq t \leq \frac{M}{F}$) and $\beta \in [0, 1]$ is the proportion of boiling noise injected.

The Signal to Noise Ratio (SNR) is calculated by the formula:

$$\text{SNR} = 10 \log_{10} \left(\frac{\sigma_b^2}{\sigma_s^2} \right)$$

where σ_b^2 is the variance of boiling noise and σ_s^2 is the variance of the signal without boiling. As both signal are divided by their standard deviation before the weighted addition (1), the SNR can be calculated as:

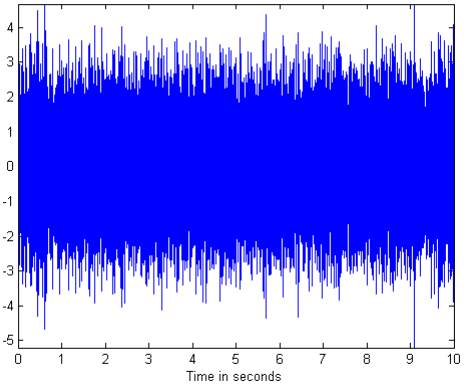
$$\text{SNR} = 20 \log_{10} \left(\frac{\beta}{1-\beta} \right)$$

An example of time domain representation of non-boiling signal and boiling signal with different SNR is provided by Figure 4. A comparison of the spectra of the non-boiling background signal of the LMFBR and the water boiling noise record is provided by Figure 5.

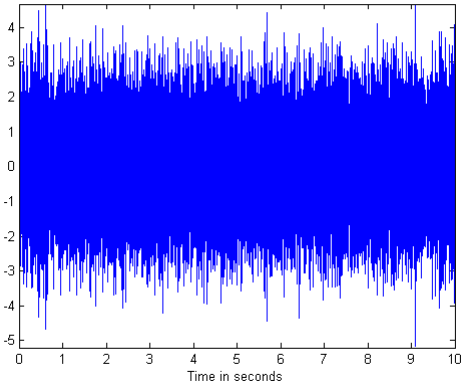
From the figures 4 and 5, it can be noticed that the LMFBR non-boiling signal has dominant components between 2 and 4 kHz while the water boiling signal has dominant components for frequencies greater than 2 kHz. Moreover, in time domain, it is difficult to distinguish graphically the non-boiling signal from the boiling signal with different SNR.

3 Presentation of the method

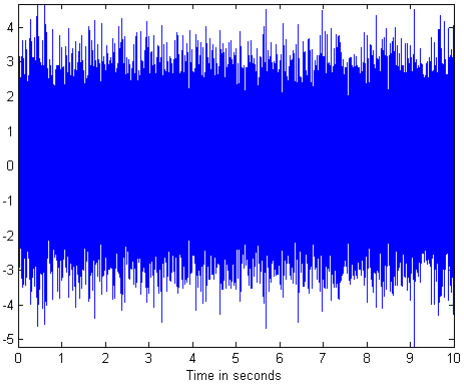
Our proposed method (see figure 6) is a supervised learning process that consists in two steps: learning and test. In the learning step, we propose to divide each acoustic background signal of the learning database into little non-overlapping time windows. A sliding window without overlapping is used for this purpose. Afterwards, the AR models corresponding to each window are estimated and two statistical techniques (multiple linear regression (LR) and Support vector machines (SVM)) are used for characterization of the current AR model class (non-boiling or boiling). The LR uses an off-line estimated normal model of background noises in no boiling condition and detects a fault if the current estimated model deviates from the normal model. The SVM starts with off-line classification of boiling and no boiling background noises models into two different classes and online supervision is done by classifying the current estimated model into one of the classes. The AR models estimation method is proposed, afterwards the classification methods are introduced.



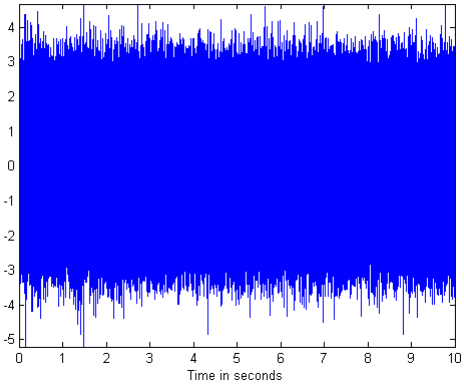
(a) Non-boiling signal in time domain



(b) SNR of -19 dB



(c) SNR of -3.5 dB



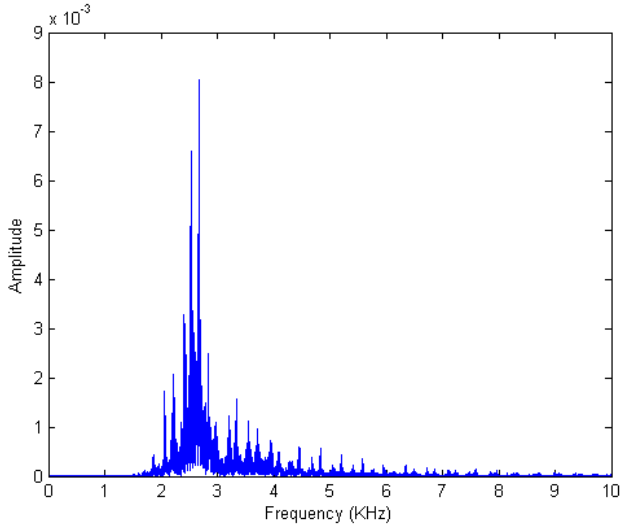
(d) SNR of 12 dB

Figure 4: Comparison of the time domain data of non-boiling signal and boiling signal for different SNR.

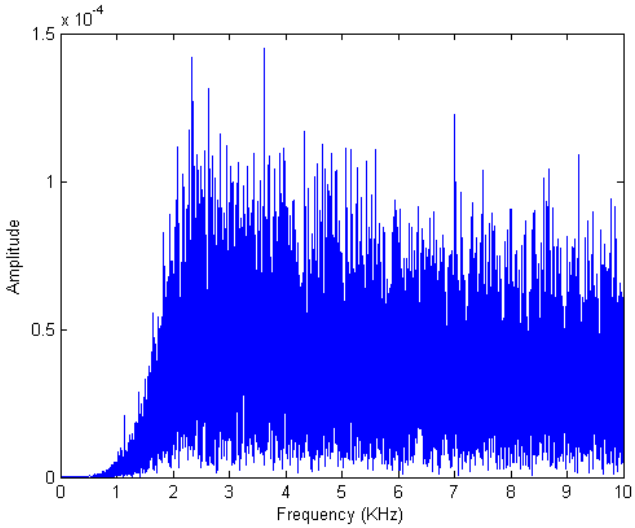
3.1 AR model estimation

Consider records of N samples by R different sensors $s_w = \left(s_w^1\, s_w^2 \ldots s_w^R\right)^T$ of the signal on time window N/F, where w is the signal number, s_w^r is the vector of N samples recorded by the sensor r $(r = 1, \cdots, R)$ and F is the sampling frequency. Assuming s_w^r is stationary, the p_r-order AR model associated to s_w^r is the vector $a_w^r = \left(a_{w,1}^r, a_{w,2}^r, \ldots, a_{w,p_r}^r\right)$ such as each sample $s_w^r[n]$ can be estimated as a linear combination of the past p_r values $(s_w^r)_{[n-p_r:n-1]} = (s_w^r[n-p_r], s_w^r[n-p_r+1], \ldots, s_w^r[n-1])^T$ as described by equation (2):

$$\widehat{s}_w^r[n] = -a_w^r\ (s_w^r)_{[n-p_r:n-1]} \tag{2}$$



(a) A non-boiling background noise spectrum



(b) A water boiling noise spectrum

Figure 5: Comparison of a power plant non-boiling spectrum with a water boiling noise spectrum.

a_w^r is calculated by minimizing the total prediction error calculated as:

$$\begin{aligned} E_w^r &= \frac{1}{N}\sum_{k=1}^{N}\left(s_w^r[k] - \widehat{s}_w^r[k]\right)^2 \\ &= \frac{1}{N}\sum_{k=1}^{N}\left(s_w^r[k] + a_s^r\left(s_w^r\right)_{[k-p_r:k-1]}\right)^2 \end{aligned}$$

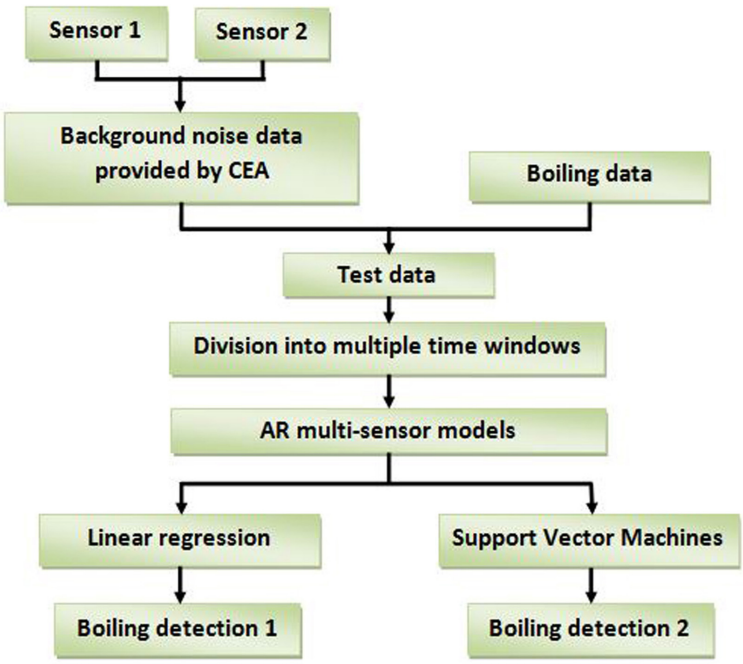


Figure 6: Summary of our proposed method

It has been proved (Makhoul, 1975; O'Shaughnessy, 1988) that a_w^r is solution of the equation (3):

$$\Phi_w^r a_w^r = -\phi_w^r \tag{3}$$

with ϕ_w^r a vector of dimension p_r defined by the formula:

$$\phi_w^r[i] = \sum_{k=i}^{N} s_w^r[k] s_w^r[k-i], \quad i = 1, \ldots, p_r \tag{4}$$

and Φ_w^r a $p_r \times p_r$ symmetrical matrix defined as:

$$(\Phi_w^r)_{i,j} = \phi_w^r[i-j], \quad 1 \leqslant i, j \leqslant p_r \tag{5}$$

The equations (3) are known as the *Yule-Walker equations*.

With the R sensors of the LMFBR, R different AR models are estimated solving Yule-Walker equations. One could combine these R models into a unique multi-sensor AR model:

$$a_w = \left(a_{w,1}^1, \cdots, a_{w,p_1}^1, a_{w,1}^2, \cdots, a_{w,p_2}^2, \ldots a_{w,1}^R, \cdots, a_{w,p_R}^R\right)^T$$

Let $p = \left(\sum_{r=1}^{R} p_r\right)$ the total number of components of the multi-sensor model. In the rest of this paper, the multi-sensor model a_w is used instead of the R one-sensor models a_w^r $(r = 1, \ldots, R)$.

3.2 The monitoring method

Once the multi-sensor AR model is obtained, the next step is to classify this latter model as boiling model or non-boiling model. Two different statistical techniques are used for this purpose: multiple Linear Regression (LR) and Support Vectors Machines (SVM). The LR monitoring method is a model based method built from a learning database of non-boiling models. It predicts the current model characteristics in non-boiling and the monitoring is made by comparing the true current model characteristics to the ones predicted. As for the SVM monitoring method, it is built from a learning database of two classes: non-boiling and boiling models. The monitoring is done by classifying the current model into one of these two classes.

3.2.1 Model based approach

We assume that, in non-boiling functioning of the LMFBR, some useful information can be extracted from the values of the environment parameters in order to characterize the multi-sensor AR models in normal functioning of the LMFBR. Let's d be the total number of environment parameters of the LMFBR taken into consideration. We therefore assume that in non-boiling functioning, any multi-sensor AR model a_w could be written as:

$$a_w = \overline{a} + Bx_w + e_w \tag{6}$$

where $\overline{a} = (\overline{a}_1, \cdots, \overline{a}_p)^T$ is an unknown mean multi-sensor model independent from w,

$$B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,d} \\ b_{2,1} & \cdots & b_{2,d} \\ \vdots & \cdots & \vdots \\ b_{p,1} & \cdots & b_{p,d} \end{pmatrix} \in \mathbb{R}^{p \times d} \tag{7}$$

is a constant unknown matrix to be determined, $x_w = (x_{w,1}, x_{w,2}, \ldots, x_{w,d})^T$ with $x_{w,j}$ ($j = 1, \ldots, d$) the value of the $j-$th environment parameter (also called *descriptor*) for the model a_w and
$e_w = (e_{w,1}, e_{w,2}, \ldots, e_{w,p})^T$ is a noise vector with zero expectation and constant variance. Bx_w can be considered as the variability of a_w explained by the environment parameters and e_w the non explained part.

The unknown vector $\overline{a}$ and matrix B can be estimated on a learning set $\{a_1, \ldots, a_L\}$ of non-boiling functioning multi-sensor models for which the value of the operating parameters are known. Indeed, equation (6) is equivalent to the following one:

$$a_w^T = \overline{a}^{\,T} + x_w^T B^T + e_w^T \tag{8}$$

Equation (8) can be written for the whole learning set as:

$$\begin{bmatrix} a_1^T \\ \vdots \\ a_L^T \end{bmatrix} = \begin{bmatrix} \overline{a}^{\,T} \\ \vdots \\ \overline{a}^{\,T} \end{bmatrix} + \begin{bmatrix} x_1^T \\ \vdots \\ x_L^T \end{bmatrix} B^T + \begin{bmatrix} e_1^T \\ \vdots \\ e_L^T \end{bmatrix}$$

or simply:

$$\begin{aligned} A &= \overline{A} + XB^T + E \\ &= 1_L\, \overline{a} + +XB^T + E \end{aligned}$$

with $1_L = (1, \cdots, 1)^T$ a vector of size L which components all equal 1. Finally,

$$A = YD^T + E \tag{9}$$

with $D = [\overline{a}\,|B]$ is the concatenation of vector $\overline{a}$ and B matrix and $Y = [1_L\,|X]$ is the concatenation of vector 1_L and X matrix.

In equation (9), A and Y (called the *learning design matrix*) are known but E is unknown. The least-squares estimator $\widehat{D}$ of D (Montgomery et al., 2012; Saporta, 2006) is given as:

$$\widehat{D} = A^T Y \left(Y^T Y\right)^{-1} \tag{10}$$

Once the estimator $\widehat{D}$ of D is calculated, an estimation of A can be calculated as: $\widehat{A} = Y\widehat{D}^T$. For any multi-sensor AR model a_w for which the descriptors values $x_w = (x_{w,1}, x_{w,2}, \ldots, x_{w,d})^T$ are known, a prevision of a_w in non-boiling functioning can be calculated as:

$$\widehat{a_w} = \widehat{D} y_w \tag{11}$$

where $y_w = [1\,|\,x_w^T]^T$.

For a given $\alpha \in [0, 1]$, a $(1-\alpha)-$prevision interval on the $i-$th component of a_w can be calculated as described in (Saporta, 2006):

$$(\widehat{D}_{i,\bullet})y_w \pm \gamma\widehat{\sigma}_i\sqrt{1 + y_w^T\left(Y^T Y\right)^{-1} y_w} \tag{12}$$

with $\widehat{D}_{i,\bullet}$ a $1-$row matrix whose components are the $i-$th row of $\widehat{D}$, γ the $\left(1 - \frac{\alpha}{2}\right)-$quantile of the Student distribution with $L - d - 1$ degrees of freedom and $\hat{\sigma_k}$ calculated as:

$$\hat{\sigma}_i = \sqrt{\frac{\|A_{i,\bullet} - \widehat{A}_{i,\bullet}\|^2}{L - d - 1}}$$

with $A_{i,\bullet}$ the $i-$th row of A.

Validity of the model Some precautions must be taken as far as LR is concerned:

1. The matrix $Y^T Y$ (equation (10)) must be invertible i.e. its determinant must not equal or near zero. A useful criteria to check that is the 2-norm condition number (the ratio of the largest eigenvalue of $Y^T Y$ to the smallest). It is well known that large condition numbers indicate a nearly singular matrix.

2. **Steadiness of the D matrix**: It is important to check that none of the models of the learning set significantly influences the values of the components of the $\hat{D}$ matrix. In other words, it must be checked that if one learning model is withdrawn from the learning set, the components of $\hat{D}$ will not completely change. The *Cook distance* (Saporta, 2006) can be used for this purpose. To check that an observed multi-sensor model a_w of the learning set, corresponding to the $w-$th row of Y, does not influence significantly the value of the $i-$th row of $\hat{D}$, $\widehat{D}_{i,\bullet}$, Cook distance of the model a_w with regards to $\widehat{D}_{i,\bullet}$ is calculated as:

$$Z_{w,i} = \frac{\left(\widehat{D}_{i,\bullet} - \widehat{D}_{i,\bullet}^{(w)}\right)\left(\mathbf{Y}^T\mathbf{Y}\right)\left(\widehat{D}_{i,\bullet} - \widehat{D}_{i,\bullet}^{(w)}\right)^T}{(d+1)\widehat{\sigma}_i^2}$$

where $\widehat{D}^{(w)}$ is the estimate D after withdrawal of the model a_w from the learning set. $Z_{w,i} \geq 1$ means that a_w is an outlier i.e. it has a significant influence on the values of the coefficients $\hat{D}$. This criteria also will be used to check the validity of the model.

The detection process The $\widehat{D}$ matrix and the $\widehat{\sigma}_i$'s $(i = 1, \cdots, p)$ are now known. For any time window signal s_w, the corresponding multi-sensor AR model a_w will be classified into boiling or non-boiling models. If the values of the environment variables are known, one could calculate the prevision intervals for each of the p components of a_w. The main point of the detection is the following: it is assumed that in normal functioning of the reactor, the true values of the components should be in their prevision interval with $100 * (1 - \alpha)\%$ confidence ($\alpha \in [0, 1]$ the risk associated to the prevision interval). The multi-sensor model a_w can be detected as a boiling model if the number of its components out of their prevision interval exceeds a detection threshold MNOC (Maximum Number of Outliers Coefficients) (an integer between 0 and p). The detection algorithm is given by algorithm 1.

Algorithm 1 Diagnosis with AR-LR

Inputs: the $\hat{B}$ matrix, the $\hat{\sigma}$'s $(i = 1, \cdots, p)$, the learning design matrix Y, the current data s_w collected on a sliding window, a threshold MNOC $\in \{0, \ldots, p\}$ and $\alpha \in [0, 1]$.

1. Calculate the AR models corresponding to s_w by using (3).
2. Calculate the multi-sensor model a_w.
3. Calculate the $(1 - \alpha)-$prevision interval for each component of a_w using (12).
4. Calculate NOC(a_w) like the Number of Outlier Coefficients (i.e. are out of their prevision interval)
5. The multi-sensor model a_w is defective if NOC$(a_w) \geqslant$ MNOC and normal otherwise.

3.2.2 Support Vectors Machines (SVM)

In the AR-LR method, the classification of the multi-sensor models into boiling and non-boiling models is based on the values of their components and those of the environment variables. For a given SNR, SVM is used to do the same classification. The SVM method aims to find a maximal margin hyperplane separating the two classes: boiling and non-boiling models (see figure 7).

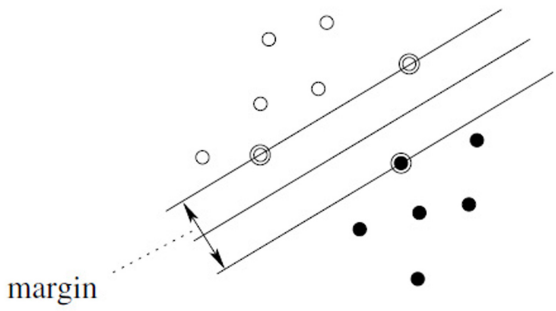


Figure 7: Illustration of the principle of SVM for classification into two classes (Ralaivola, 2010)

Let $\{a_1, a_2, \ldots, a_L\}$ the learning set of multi-sensor AR models. Unlike the LR, this learning set is now composed of both boiling and non-boiling AR models (classes are {non-boiling,

boiling} or {+1 , -1}). Let K a gaussian kernel calculated for two multi-sensor models $a_w = (a_{w,1}, \cdots, a_{w,p})^T$ and $a_v = (a_{v,1}, \cdots, a_{v,p})^T$ with the formula:

$$K(a_w, a_v) = \exp\left(-\frac{1}{2\rho^2}\|a_w - a_v\|^2\right) \tag{13}$$

where $\rho > 0$ is the *bandwidth* of K.

We define η as a function of multi-sensor models such as $\eta(a_w)$, $w = 1, \ldots, L$, equals +1 if a_w is a non-boiling multi-sensor model and -1 otherwise and α^* the vector of dimension L solution of the following optimization problem (see (Flecther, 2009) for more details):

$$\begin{cases} \max\limits_{\alpha} \sum\limits_{w=1}^{L} \alpha_w - \frac{1}{2} \sum\limits_{w,v=1}^{L} \alpha_w \alpha_v \eta(a_w)\eta(a_v)K(a_w, a_v) \\ \\ \text{subject to} \\ \sum\limits_{w=1}^{L} \alpha_w \eta(a_w) = 0 \\ 0 \leq \alpha_w \leq C, \quad w = 1, \ldots, L \end{cases} \tag{14}$$

where C is a regularization constant for outliers. Let b^* calculated with the following formula:

$$b^* = \frac{1}{L} \sum_{w=1}^{L} \left(\eta(a_w) - \sum_{v=1}^{L} \alpha_v^* \eta(a_v) K(a_w, a_v) \right)$$

Detection process The diagnosis with AR and SVM (AR-SVM) can be summarized as:

Algorithm 2 Diagnosis with AR-SVM

Input: a time window signal s_w

1. Calculate the AR models corresponding to s_w using (3).
2. Calculate the multi-sensor model a_w associated to s_w.
3. Calculate the feature $f(a_w) = \sum\limits_{v=1}^{L} \alpha_v^* \eta(a_v) K(a_v, a_w) + b^*$
4. a_w is a boiling model if $f(a_w) < 0$ and is a non-boiling model otherwise.

In our work, the SVM method was implemented using the MATLAB toolbox developed in (Loosli et al., 2004).

4 Results on the experimental data

4.1 Determination of the appropriate AR order and time windows duration

One of the main points in the AR process is to choose two correct orders p_1 and p_2 for the AR on the sensors 1 and 2 and also a correct duration l for the windows.

In order to determine a correct value for the AR order p_1, we first setted l to 20 ms and then looked for the value of p_1 that minimizes the mean value of Akaike's Information Criterion for Finite samples (AICF) (Mahmood, 2007). For a given signal divided into W time windows signals $s_w^1 \in \mathbb{R}^{N\times 1}$ $(w = 1, \ldots, W)$, the mean AICF is calculated as follows:

$$AICF(p_1) = \frac{1}{W}\sum_{w=1}^{W}\left(\ln\left(\frac{1}{N-p_1}\sum_{k=p_1+1}^{L}\left(s_w^1[k] + (a_{s_w}^1)^T(s_w^1)_{[k-p_r,k-1]}\right)^2\right)\right) + \frac{2p_1}{L-2p_1} \quad (15)$$

An example of variation of the mean AICF for values of p_1 varying from 1 upto 100 is shown on the figure 8.

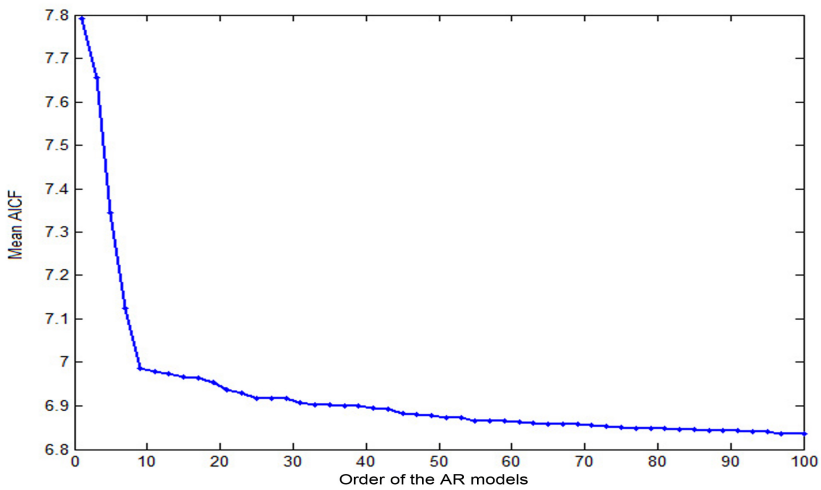


Figure 8: Example of variation of the mean AICF computed on 1000 windows of 20 ms according to AR order p_1 for the sensor 1

It can be seen on the figure 8 that the mean AICF decreases as p_1 increases. Particularly when p_1 varies from 1 to 10, the decrease is very quick. So any AR order greater than 10 can be used for the sensor 1. Here we have decided to take $p_1 = 12$. This value was chosen because no appreciable reduction of the AICF values are observed for higher model orders. We also set the AR order on the sensor 2 to $p_2 = 12$ which implies that the total number of components for any multi-sensor model is $p = 24$.

In order to apply the algorithm for real-time detection, it is necessary that the size l of the sliding window should be between 10 and 200 milliseconds. The figure 9 shows an example of variation of the AICF for different values of l.

It can be seen that the AICF decreases as l increases. Any l greater than 60 could be used and we have decided to take $l = 100$ milliseconds for both sensors.

The AR model estimation assumes that the background signal is stationary. In this work, there is not only one AR model estimated but there are many models estimated each on sliding windows of 100 ms duration. And the assumption of the background noise to be stationary is made on each of these background noises of 100 ms. On a 100 ms time interval it is possible to assume that the environment conditions of the power plant do not change and also that the system is stationary. However, on longer time intervals these conditions can change.

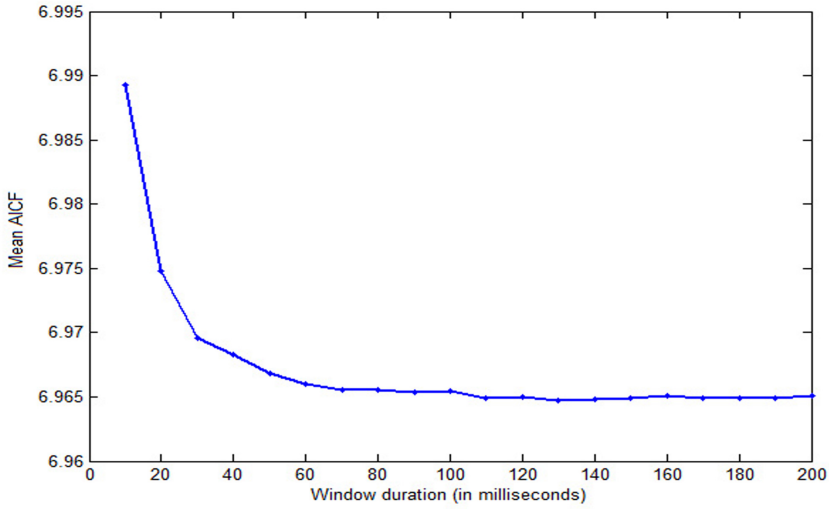


Figure 9: Example of evolution of the AICF for different values of the duration l of windows

4.2 Boiling Detection based on AR and LR

4.2.1 Validity of the model

In the study, seven operating parameters are considered ($d = 7$) that leads to a value of $\det\left(X^T X\right)$ greater than 10^{54}. Unfortunately, the $X^T X$ matrix is ill-conditioned since its 2-norm condition number is 5.0835×10^8 which is very great. An estimate of the correlation matrix of the environment parameters is given by:

$$\Sigma = \begin{bmatrix} 1.00 & & & & & & \\ 0.66 & 1.00 & & & & & \\ 0.90 & 0.92 & 1.00 & & & & \\ 0.72 & 0.12 & 0.44 & 1.00 & & & \\ 0.84 & 0.22 & 0.57 & 0.75 & 1.00 & & \\ 0.99 & 0.67 & 0.90 & 0.67 & 0.84 & 1.00 & \\ 0.78 & 0.76 & 0.85 & 0.46 & 0.55 & 0.78 & 1.00 \end{bmatrix}$$

where the environment parameters are in the order of the table 1. It can be seen that there are many almost perfect correlations between the environment parameters. In order to cancel the correlation, we perform a Principal Components Analysis (PCA) which is a dimension reduction technique consisting in finding linear combinations of the environment parameters such as the maximum variance possible is retained (see (Saporta, 2006; Dunteman, 1989)). As the units of the environment parameters are different, a correlation PCA is performed. We have chosen the first two components of the PCA for 92% of variance explained. Now we use these first two components as descriptors and we get about 5.5 as 2-norm condition number. This latter is a more acceptable value.

The maximum value of the Cook distances for each of the learning multi-sensor AR models is lower than 0.0022. That implies that none of the models of the learning dataset does influence significantly the values of the LR coefficients which means that our LR model is valid.

4.2.2 Example of Boiling detection

Detection with AR-LR algorithm (algorithm 1) is performed on one single signal of duration 20 seconds recorded by sensors 1 and 2. Film boiling (stage four of boiling) is injected at $t = 7$ seconds. The SNR is 3.5 dB. The output signal is divided into 200 windows signals of 100 milliseconds. Then the multi-sensor AR models corresponding to each window are calculated (i.e. twelve AR coefficients are evaluated for each time window signal and for each of the two sensors). For one of the 24 components of the sliding window AR multi-sensor model, say the $i-th$ component, the consecutive values for the 200 windows can be represented by a vector a_i of 200 values. Figure 10 shows the variation (the plot of the 200 consecutive values) of each of the AR coefficients in time domain for the selected signal.

From the figure 10 it is noticed that, most of the AR coefficients (18 out of 24) follow a sudden change in their values around $t = 7$ seconds. But their values remain in the prevision interval except for the coefficients a_{10} and a_{11} for each sensor. It confirms that the boiling phenomenon causes a disturbance to the values of some AR coefficients.

We introduce the *boiling detection state* as the main figure to be monitored for boiling detection in a signal. It is equal to 1 if the multi-sensor AR model corresponding to the sliding window is detected as boiling model and 0 otherwise. The boiling detection state corresponding to the selected signal of the figure 10 is shown on figure 11. The MNOC is equal to 2 i.e. the boiling state is equal to 1 if there are more than 2 components out of their prevision intervals and is zero otherwise. It can be seen that onset of boiling is clearly detected around $t = 7$ seconds.

4.2.3 False detection rates

False detection rate is defined as the rate of time windows detected as boiling windows when there is no boiling at all. In this study, after dividing all the records of the background noise set, we obtain 33 560 windows of 100 ms (in normal functioning) which are divided in two sets: 16 780 for learning and 16 780 for test. The false detection rates (also known as *false alarm rates*) corresponding to $\alpha = 0.1\%$ (99.9% confidence for the prevision intervals) and $\alpha = 0.01\%$ (99.99% of confidence) are given in the tables 2 and 3.

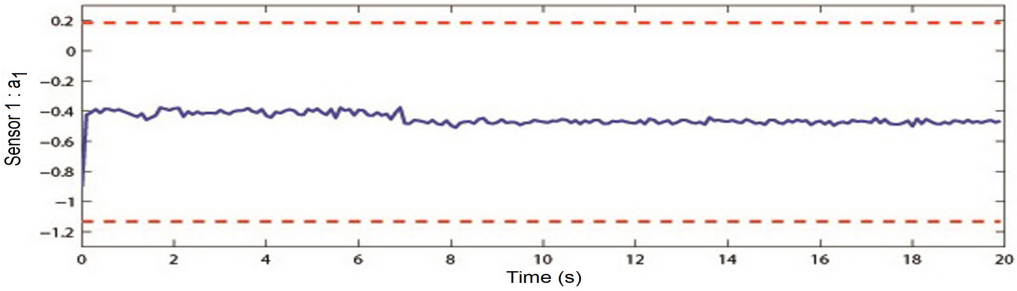
Table 2: False detection rates (in %) on the 16 780 test windows with a 99.9% confidence (risk $\alpha = 0.1\%$) for the prevision intervals.

MNOC	5	4	3	2	1	0
Rates	0.01	0.02	0.08	0.86	5.77	8.97

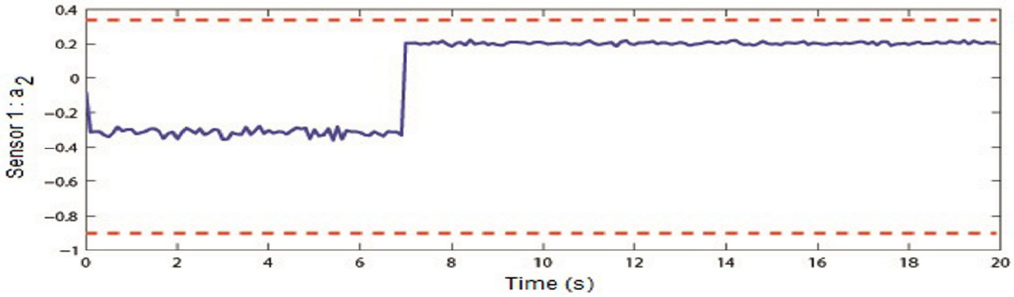
Table 3: False detection rates (in %) on the 16 780 test windows with a 99.99% confidence (risk $\alpha = 0.01\%$) for the prevision intervals

MNOC	5	4	3	2	1	0
Rates	0.00	0.00	0.00	0.01	0.44	4.41

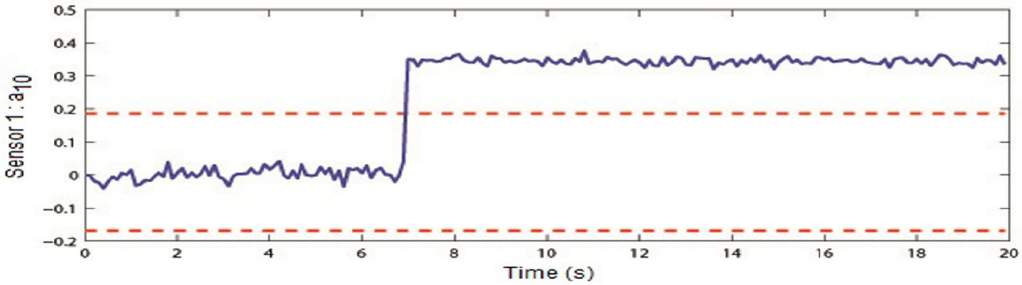
The false detection rates increases as the MNOC decreases. Indeed, MNOC can be seen as



(a) Example of unchanged evolution: coefficients a_1 (sensor 1) and a_8 (sensor 2).



(b) Example of coefficients whose values are clearly disturbed around 7 seconds but remain in their prevision intervals: coefficients a_2 to a_9 and a_{12} for sensor 1; coefficients a_1 to a_7, a_9 and a_{12} for sensor 2.



(c) Example of coefficients whose values are out of their prevision intervals: coefficients a_{10} and a_{11} for both sensors 1 and 2.

Figure 10: Time domain variation of the components of the sliding multi-sensor AR model with SNR = 3.5 dB: the lower and the upper bounds of the prevision intervals are in dotted line, the true values are in continued line.

a level of freedom given to the multi-sensors models. For both values of the risk α, for MNOC greater than 5, the false detection rate is zero. It means that all the 16 780 models of the test set have at least 19 components present in their prevision intervals. It seems that the value MNOC = 0 (no outlier coefficient allowed) is not too good because of the high rates of false alarms compared to the rates of the other values of MNOC. Except for MNOC= 0, the false detection rates presented in tables 2 and 3 are very low and almost zero, which means that most of the time the real values of the components of the AR models are present in their respective confidence intervals. Most importantly, it means that the models are not disturbed by a change in the operating mode as the test set contains windows with different operating modes. As expected,

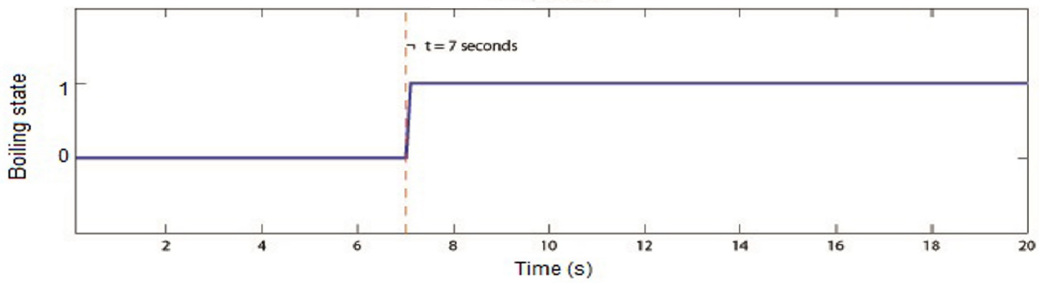


Figure 11: Boiling detection state for the selected signal containing film boiling from 7 seconds. The SNR is 3.5 dB and the MNOC is equal to 2. The boiling detection state equals 1 if boiling is detected in the sliding window and 0 otherwise.

the false alarm rate is lower when the chosen risk is lower.

4.2.4 Detection rates

We define the *detection rate* as the rate of windows detected as boiling windows when there is boiling. For a given value of SNR, boiling has been injected in each of the 33 560 normal functioning time windows from $t = 0$. Afterwards, algorithm 1 was applied on each of the 33 560 windows. The detection rates are almost the same for each type of boiling. Those corresponding to film boiling are presented in tables 4 (for a risk $\alpha = 0.1\%$) and 5 (for a risk $\alpha = 0.1\%$).

Table 4: Detection rates (in %) for 99.9% confidence (risk $\alpha = 0.1\%$) for the prevision intervals and for film boiling.

SNR	MNOC								
	8	7	6	5	4	3	2	1	0
-19	1.6	3.8	4.02	4.71	29.9	54.4	78.4	83.0	88.7
-12	0.1	9.5	35.8	52.7	73.8	84.9	88.7	92.7	95.5
-7	0.4	20.7	52.8	71.4	80.0	91.7	93.0	97.9	99.9
-3.5	0.6	16.3	63.0	74.7	83.3	92.7	94.7	99.5	100.0
0	0.1	4.8	62.3	74.6	82.5	92.9	97.6	99.9	100.0
3.5	0.0	0.1	48.0	68.7	78.9	95.0	99.4	99.9	100.0
7	0.0	0.0	1.5	16.2	77.6	97.7	99.9	100.0	100.0
12	0.0	0.0	0.0	10.7	59.0	98.1	100.0	100.0	100.0
19	0.0	0.0	0.0	9.5	76.2	99.9	100.0	100.0	100.0
40	0.0	0.0	0.0	14.5	86.2	100.0	100.0	100.0	100.0

The detection rates increases as MNOC decreases. For MNOC = 1, 2 or 3, the detection rate increases as the SNR increases and is always very close to 100%. Which means good boiling detection. As expected, these detection rates corresponding to $\alpha = 0.01\%$ are much lower than those of $\alpha = 0.1\%$ because when the risk α decreases, the amplitude of the confidence intervals increases and there is much less detection of outliers coefficients.

One critical issue is to find an optimal value for MNOC. Indeed, a low MNOC (say 0 or 1) leads to a high false alarm rates and a high detection rate while a high MNOC leads to an almost

Table 5: Detection rates (in %) for 99.99% confidence ($\alpha = 0.01\%$) for the prevision intervals and for film boiling. The SNR is given in dB.

SNR	MNOC								
	8	7	6	5	4	3	2	1	0
-19	0.0	0.0	0.2	1.9	26.6	42.2	68.2	74.6	82.7
-12	0.0	0.0	1.7	9.6	40.0	78.1	84.4	92.6	93.1
-7	0.0	0.0	3.3	52.2	56.0	87.3	91.3	93.2	97.4
-3.5	0.0	0.0	4.6	54.5	64.7	92.5	92.7	97.5	100.0
0	0.0	0.0	0.3	53.6	64.3	92.5	94.7	99.0	100.0
3.5	0.0	0.0	0.0	14.7	44.2	92.6	97.3	100.0	100.0
7	0.0	0.0	0.0	0.2	2.1	91.7	96.7	100.0	100.0
12	0.0	0.0	0.0	0.0	0.1	88.5	99.3	100.0	100.0
19	0.0	0.0	0.0	0.0	0.4	89.4	99.9	100.0	100.0
40	0.0	0.0	0.0	0.0	35.5	99.8	100.0	100.0	100.0

zero value for the false alarm rates but also to a low detection rate. To choose the optimal value of MNOC, we look for the MNOC between 0 and 5 that maximizes detection rate and minimizes false alarms rate. This is equivalent to maximizing the *feature* function: *detection rate* − *false alarms rate*. We calculate the mean detection rates per column of tables 4 and 5 and we deduce respectively the false alarms rates given by tables 2 and 3.

Table 6: Values of the feature function (*detection rate* − *false alarms rate*) for MNOC between 0 and 5 and the risk $\alpha = 0.1\%$.

MNOC	5	4	3	2	1	0
Feature function	39.8	72.8	90.7	94.4	91.5	89.5

Table 7: Values of the feature function (*detection rate* − *false alarms rate*) for MNOC between 0 and 5 and the risk $\alpha = 0.01\%$.

MNOC	5	4	3	2	1	0
Feature function	18.7	33.4	85.5	92.5	95.2	92.9

We propose as optimal values: MNOC = 2 with the risk $\alpha = 0.1\%$ and MNOC = 1 with the risk $\alpha = 0.01\%$.

4.3 Boiling Detection by AR and SVM

For different values of the SNR, we have applied the algorithm 2 (AR-SVM). The non-boiling set is composed of 33 560 non-boiling multi-sensors models. For a given value of the SNR and after injection of boiling from $t = 0$ second in these non-boiling windows, we finally get a set of

67 120 multi-sensor models. The set is then divided into learning set (50%) and test set (50%). The results of the classification on the test set are almost the same for each of the four types of boiling. Those corresponding to film boiling are presented in table 8. *True Negative* (TN) corresponds to the models classified as non-boiling when there is no boiling at all, *False Negative* (FN) corresponds to the models classified as non-boiling when there is boiling actually, *False Positive* (FP) corresponds to the models classified as boiling when there is no boiling and *True Positive* (TP) corresponds to the models classified as boiling when there is really boiling. The *False Alarm Rate* (FAR) is calculated as the ratio of FP to the total number of non-boiling models in the test set (i.e. 16780) and the *Detection Rate* is calculated as the ratio of TP to the total number of boiling models in the test set (i.e. 16780). *Error* is calculated as the ratio of misclassified models to the total number of models (33560).

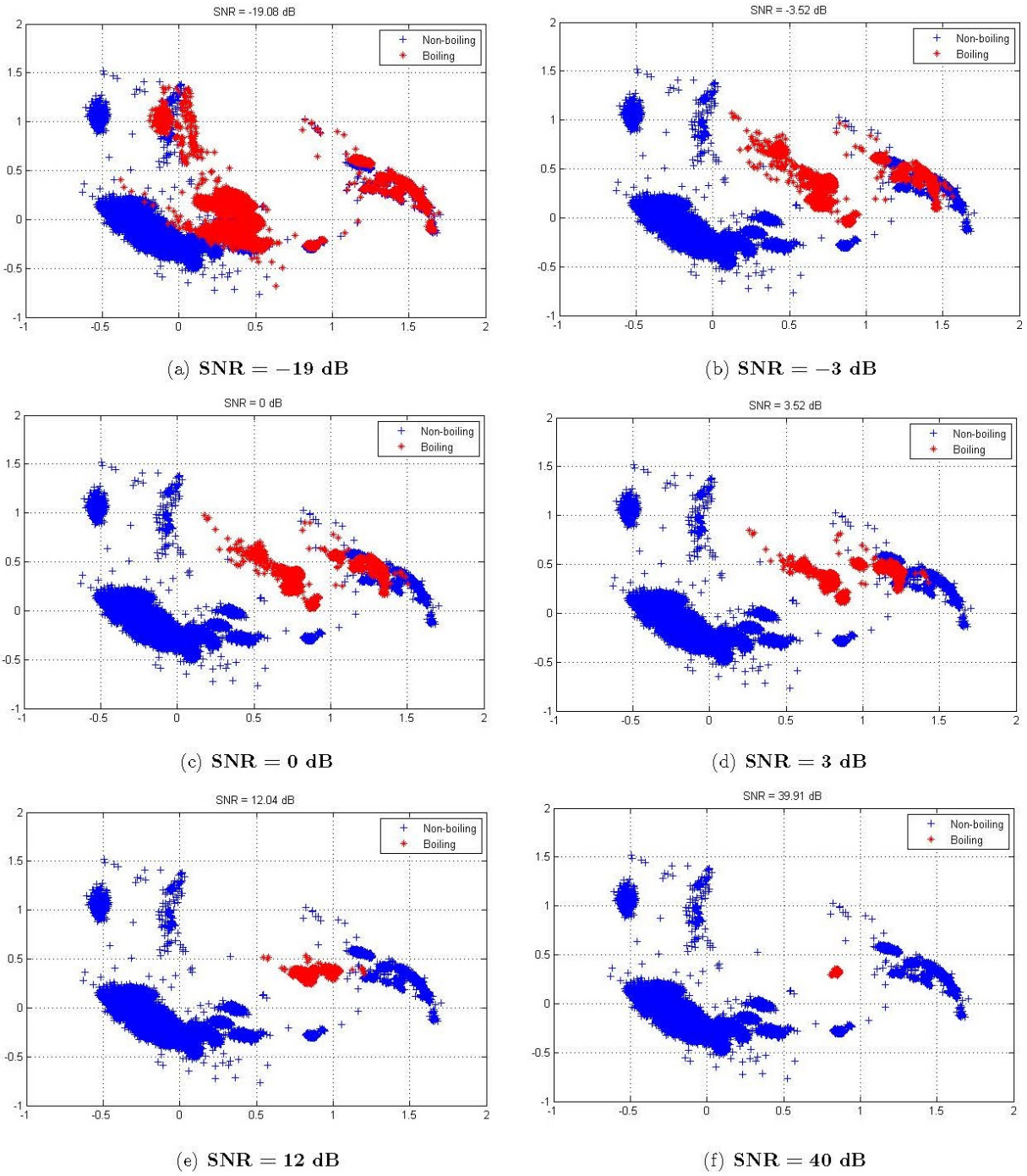
Table 8: Results of the SVM classifier on the test data. The boiling models correspond to injection of film boiling. The SNR is given in dB.

SNR	TN	FN	FP	TP	Error	FAR	DR
-19	16773	12	7	16768	0.06%	0.0004%	99.93 %
-12	16779	3	1	16777	0.01%	0.00005%	99.98%
-7	16779	0	1	16780	0.003%	0.003%	100%
-3.5	16780	0	0	16780	0%	0%	100%
0	16780	0	0	16780	0%	0%	100%
3.5	16780	0	0	16780	0%	0%	100%
7	16780	0	0	16780	0%	0%	100%
12	16780	0	0	16780	0%	0%	100%
19	16780	0	0	16780	0%	0%	100%
40	16780	0	0	16780	0%	0%	100%

The multi-sensor models can be considered as elements of $\mathbb{R}^{24}$. To view them, we performed a Covariance Principal Components Analysis (PCA) on the 33 560 multi-sensor models in non-boiling conditions afterwards the projections of the 33 560 boiling multi-sensor models in the principal plan of the PCA (the first two components) are calculated and shown in the figure 12 for increasing values of the SNR. It can be seen that with the increase of the SNR, the boiling multi-sensor models cluster at one point which means a better classification and thus a better boiling detection.

5 Conclusion

In this work, we have presented new approaches for acoustic monitoring of a liquid metal fast breeder reactor as far as the sodium boiling detection is concerned. The proposed method is based on the estimation of AR model parameters on a sliding window. Afterwards, the boiling detection is done by LR and SVM methods. The LR uses an off-line estimated normal model of background noises in non-boiling condition and detects a fault if the current estimated model deviates from the normal model. The SVM starts with off-line classification of boiling and non-boiling background noises models into two different classes and online supervision is done by classifying the current estimated model into one of the classes. In order to avoid wrong detections due to operating modes changing, their parameters are taken into account in the



(a) **SNR = −19 dB**

(b) **SNR = −3 dB**

(c) **SNR = 0 dB**

(d) **SNR = 3 dB**

(e) **SNR = 12 dB**

(f) **SNR = 40 dB**

Figure 12: Projection of boiling multi-sensor models in the principal plan of PCA performed on non-boiling models only: non-boiling models are in blue and boiling models are in red. The first two principal components explain 73.68% of the total variance.

detection method. The results of the proposed method applied to the test data prove that boiling phenomenon introduces disturbance in the values of the AR models (or coefficients) than

can be clearly detected. The high detection rates as well as the low false alarms rates obtained on the test data show that our proposed approach is successful even for negative values of the Signal to Noise Ratio. Most importantly, the proposed method is not sensible to the change of LMFBR operating mode. In order to obtain more representative acoustic signals, work for characterization of the acoustic signature of the liquid sodium boiling noise corresponding to core accidents is in progress. Further work will also consider definition of a parametric distance to reduce the number of AR coefficients to be monitored in the AR combined to linear regression detection method.

Acknowledgement

This work was supported by GIS-3SGS COSMOS: Contrôle et Observation de Systèmes à Modes multiples de fonctionnement Surs. The paper is made in the framework of the *DA Coeur* (Acoustic Detection of Core Accident) project initiated by the CEA (Commissariat à l'Energie Atomique et aux énergies alternatives, France). The authors heartily thank its whole staff who accepted to give the experimental data used in this work and especially Christian Lhuillier who made the recording of the data at Phenix nuclear power plant.

References

Bomeluberg, H. J. (1968). An evaluation of the applicability of water model testing to liquid metal engineering problems. Technical report, Liquid Metal Engineering Center.

Carey, W. M. (1979). The acoustic background noise of a liquid metal fast breeder reactor. *Journal of the Acoustical Society of America*, 65(6):1461–1466.

Dall'Ava, D., Martin, L., and Vray, B. (2009). 35 years of operating experience of phenix npp sodium cooled fast reactor. In *Proceedings of the 17th International Conference on Nuclear Engineering (ICONE 17)*, volume 1, pages 243–254. American Society of Mechanical Engineers (ASME). Brussels, Belgium, 12-16 july, 2009.

Dunteman, G. H. (1989). *Principal Components Analysis*. Number 69 in Quantitative Applications in the Social Sciences. SAGE Publications.

Flecther, T. (2009). Support Vector Machines Explained. http://www.tristanfletcher.co.uk/SVM%20Explained.pdf. Last accessed: July 3rd, 2013.

Hayashi, K. (1997). Autoregressive techniques for acoustic detection of in-sodium water leaks. In *Acoustic signal processing for the detection of sodium boiling or sodium-water reaction in LMFRs*, number IAEA-TECOOC-946, pages 179–202. International Atomic Energy Agency (IAEA).

Hayashi, K., Shinohara, Y., Watanabe, K., and Nabeshima, K. (1996). Twice squaring method: non-linear acoustic signal processing technique for real time detection of sodium boiling. *Ann. Nucl. Energy*, 23(10):863–875.

Inujima, H., Ogino, T., Haga, K., and Kikuchi, Y. (1982). Detection of local sodium boiling in a simulated LMFBR fuel subassembly. *Progress in Nuclear Energy*, 9(0):107 – 121.

Lienhard, J. H. IV, Lienhard, J. H. V (2000). *A heat transfer textbook.* J. H. Lienhard V, Cambridge, Massachusetts, USA, 3rd edition.

Loosli, G., Canu, S., Vishwanathan, S. V. N., Smola, A., and Chattopadhyay, M. (2004). Une boîte à outils rapide et simple pour les SVM. *Presses Universitaires de Grenoble*, pages 113–128.

Mahmood, K. (2007). Finite sample AIC for autoregressive model order selection. In *2007 IEEE International Conference on Signal Processing and Communications (ICSPC 2007)*, pages 1219–1222, Dubai, United Arab Emirates.

Makhoul, J. (1975). Linear prediction : A tutorial review. *Proceedings of the IEEE*, 63(4):561 – 580.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to Linear Regression Analysis.* Wiley, 5-th edition.

O'Shaughnessy, D. (1988). Linear predictive coding. *Potentials, IEEE*, 7(1):29–32.

Paumel, K., Jeannot, J., Jeanne, T., Laffont, G., Vanderhaegen, M., and Massacret, N. (2013). R&D on Early Detection of the Total Instantaneous Blockage for 4th Generation Reactors: Inventory of non-Nuclear Methods Investigated by the CEA. In *2013 3rd International Conference on Advancements in Nuclear Instrumentation, Measurement Methods and their Applications (ANIMMA).* Marseille, France, 23-27 June 2013.

Prakash, V., Thirumalai, M., Anandaraj, M., Kumar, P. A., Ramdasu, D., Pandey, G., Padmakumar, G., Anandbabu, C., and Kalyanasundaram, P. (2011). Experimental qualification of subassembly design for Prototype Fast Breeder Reactor. *Nuclear Engineering and Design*, 241(8):3325 – 3332.

Ralaivola, L. (2010). Introduction to kernel methods. Pattern Analysis, Statistical Modelling and ComputAtional Learning (PASCAL) Bootcamp in Machine Learning, Marseille, France. this issue.

Saporta, G. (2006). *Probabilités, analyse des données et statistique.* Editions TECHNIP, 2è edition.

Vanderhaegen, M., Paumel, A., and Tourin, A. (2013). Acoustic emissions of a boiling liquid - an experimental survey in water and extrapolation to SFRs. Third International Conference on Advancements in Nuclear Instrumentation, Measurement Methods and their Applications (ANIMMA 2013). 23-27 June, Marseille, France.