



HAL
open science

An Original Method for Detection of a Known Anomaly: Application to the Detection of Sudden Coolant Blockage

Rémi Cogranne

► **To cite this version:**

Rémi Cogranne. An Original Method for Detection of a Known Anomaly: Application to the Detection of Sudden Coolant Blockage. ACM 10th International Conference on Multimedia Systems and Signal Processing, May 2025, Fukui, Japan. ⟨hal-04895600⟩

HAL Id: hal-04895600

<https://hal.science/hal-04895600v1>

Submitted on 18 Jan 2025

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.



HAL Authorization

An Original Method for Detection of a Known Anomaly: Application to the Detection of Sudden Coolant Blockage

Rémi Cogranne
Troyes University of Technology
Troyes, France
remi.cogranne@utt.fr

ABSTRACT

Ensuring the safe operation of critical systems is of the highest importance, particularly when the consequences of a malfunction can be catastrophic. The present paper focuses on developing a reliable and rapid detection method for a critical event in nuclear power plants: the total and instantaneous blockage of coolant flow. Our approach involves modelling the temperature of fuel rod assemblies, taking into account past temperature readings to improve accuracy. By distinguishing between normal and abnormal temperature patterns, we can identify potential blockages early on. The proposed model and associated statistical test are versatile and can be applied to various nuclear reactors and other anomaly detection scenarios. Our results, based on real-world temperature data from the Superphénix power station, demonstrate the accuracy of the proposed model and the relevance of the detection procedure.

CCS CONCEPTS

• **Computing methodologies** → **Information extraction**; *Learning linear models*; *Probabilistic reasoning*; Anomaly detection; Control methods; • **Hardware** → **Signal processing systems**; *Safety critical systems*; Temperature control; • **Mathematics of computing** → **Stochastic processes**.

KEYWORDS

Multidimensional Signal Processing, Sequential Detection, Coolant Blockage, Temperature Control, Real-Time, Monitoring

ACM Reference Format:

Rémi Cogranne. 2025. An Original Method for Detection of a Known Anomaly: Application to the Detection of Sudden Coolant Blockage. In *Proceedings of 10th ACM International Conference on Multimedia Systems and Signal Processing (ICMSSP 2025)* (ACM ICMSSP'25). ACM, New York, NY, USA, 8 pages. <https://doi.org/XXXXXXXX.XXXXXXX>

1 INTRODUCTION

Monitoring and supervision critical infrastructure is a crucial task. Among these infrastructures, nuclear power plants are particularly important and sensitive due to the significant consequences of potential incidents.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM ICMSSP'25, May 9 – 11, 2025, Fukui, Japan

© 2025 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXXX.XXXXXXX>

As new reactor types are being developed, it is essential to investigate how to monitor these complex systems to detect potential malfunctions. The Sodium-cooled Fast Reactor (SFR) is one of the six models selected by the Generation IV International Forum (GIF), which aims to provide recommendations, conduct further studies, and assess technologies for future nuclear power plants [19]. Despite careful development, security systems may still be insufficient when rare events occur [2]. Incidents related to core cooling are particularly concerning, as they can have severe consequences. One of the most feared incidents is the Total and Instantaneous Blockage (TIB), which, although unlikely, can cause significant damage if the detection of such an event is missed. This paper addresses the detection of TIB.

Monitoring the temperature of the nuclear reactor core in real time enables the detection of overheating and other types of simple incidents. However, some incidents may be more challenging to detect. In this context, it is vital to develop a highly reliable detector. This means that the detector should have statistically established properties that guarantee a low rate of false alarms and minimize the probability of missed detections.

1.1 State of the Art

Detecting thermal incidents is a complex problem that falls under the broader category of monitoring complex systems. Change detection, which analyses data as it is received, is a central issue in this field. While change detection has been widely studied in various applications, such as finance, econometrics, and manufacturing processes, existing approaches often lack known probabilities for false alarms and missed detections.

The first sequential detection method, developed by Wald, focused on cases where observations follow one of two models from the beginning. This led to the development of the Sequential Probability Ratio Test (SPRT), which has been shown to have several optimal properties. However, the SPRT is not suitable for cases where observations can change after a certain point.

The CUSUM method was later proposed for change-point detection, which "resets" the SPRT to reduce detection delays. CUSUM has been extensively studied and has been shown to be optimal for change-point detection in several cases.

The problem of detecting Total and Instantaneous Blockage (TIB) is an example of such events difficult to identify. Indeed, unlike traditional change-point detection methods, which assume a sudden change in distribution parameters, TIB detection involves a complex scenario where the temperature of the central rod decreases while neighbouring rods continue to increase.

Therefore the problem of detection of Total and Instantaneous Blockage requires a combination of multidimensional signal processing, to analyse the temperature of all rods in time, along with a sequential detection methodology, to address the problem of quickest detection.

The problem of detecting Total and Instantaneous Blockage (TIB) in nuclear reactors has received limited attention, particularly for Sodium-cooled Fast Reactors (SFR), a new technology still under development [15]. Previous studies have proposed various approaches, but they have limitations. For example, a dynamic ARMAX model was proposed to describe reactor core temperature [13], but it does not provide a detection method and may not be suitable for TIB detection. Another approach used a set of rules to compare neighbouring rod temperatures [7], but its reliability is questionable due to the lack of statistical analysis. More recently, a neural network-based approach was proposed [10], but its performance is also uncertain. Furthermore, supervised learning-based detection schemes may struggle to analyse rare events and require retraining for different reactor technologies. In contrast, this paper presents a new approach for reliable, accurate and the quickest detection of TIB.

1.2 Contribution and Organization of this Paper

The present paper proposes three main contributions for the detection of a known anomaly with application to the problem of Total and Instantaneous Blockage (TIB) of coolant in nuclear reactors. Firstly, we propose a new adaptive linear model that accurately captures the complex temperature behaviour of reactor rods under normal operating conditions. This model is unique in that it uses previous temperature measurements to refine its predictions, allowing for a more precise representation of the reactor's thermal dynamics.

Secondly, our approach is versatile and can be applied to a wide range of reactors with different experimental settings, without requiring any prior training or learning phase. This makes it a valuable tool for the nuclear industry, where reactors can vary significantly in design and operation.

Thirdly, our adaptive linear model enables the design of a reliable statistical test for TIB detection. By combining our model with a simple yet efficient method for the detection of a known change, we can guarantee a high probability of TIB detection within a specified time frame, while also controlling the rate of false alarms. This ensures a reliable and efficient detection system for TIB events.

Our proposed methodology enables reliable TIB detection, which is verified through numerous numerical simulations on real data.

The present paper is organized as follows. First Section 2 explains the problem of Total and instantaneous blockage of the coolant and formalizes the underlying detection problem. Then Section 3 present in detail the model we proposed to represent the non-anomalous temperature. The statistical test for the detection of Total and Instantaneous Blockage using this model is presented in section 4 and its statistical performance are analytically established. In addition, this section also presents the sequential detection method used in the present paper to minimize the average detection delay. Then the Section 5 presents and discusses a wide range of numerical results, which show the relevance of the

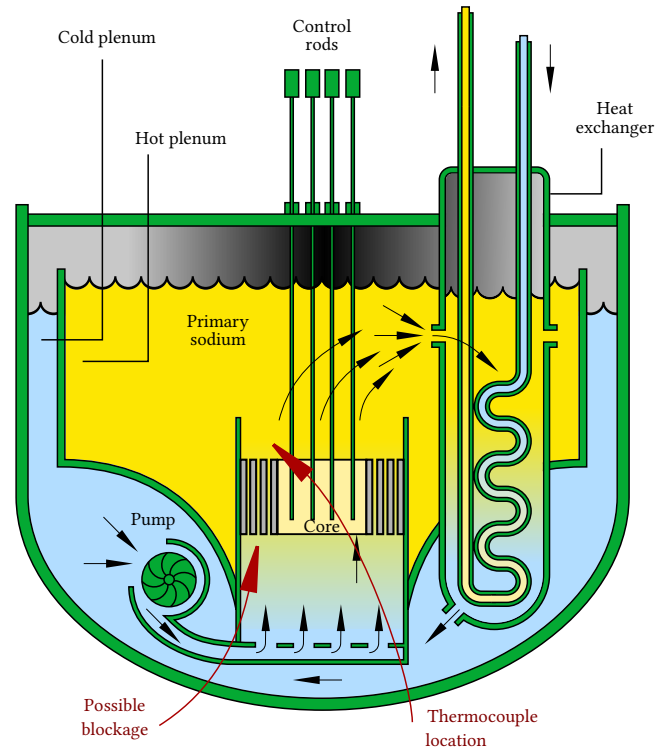


Figure 1: Illustration of the overall design of an SFR nuclear reactor, here using a pool architecture emphasizing the location of thermocouples with respect to the possible blockage. Source: based on an image in the public domain from www.Gen-4.org.

proposed original model and the accuracy of the ensuing statistical test. Section 6 concludes the paper and draws possible future works.

2 KNOWN ANOMALY DETECTION: STATEMENT OF THE PROBLEM

This section provides a detailed explanation of the TIB event that our detection method aims to identify. We also highlight the challenges associated with this task and introduce the notation used throughout the paper. First, we describe in Section 2.1 the problem of total and instantaneous blockage along with its model on the temperature of nuclear core rods. We then formally introduced in Section 2.2 the statistical problem of detection of a known anomaly, in our case the blockage.

2.1 Total and Instantaneous Blockage of Coolant

A SFR nuclear reactor consists of a core with rods and a coolant, which ensures cooling and heat exchange with an alternator that produces electricity. The rods are arranged in a honeycomb-like structure, with each rod sheltered in a steel pipe where the coolant flows. The temperature of the assembly of rods is monitored carefully to detect overheating. However, it is not possible to directly measure the temperature of the rods, so the temperature of the output coolant is measured instead.

A blockage occurs when the coolant flow is completely and suddenly stopped, which can lead to a major accident [5]. The most famous case of a Total and Instantaneous Blockage (TIB) occurred on October 5, 1966, at the FERMI-1 station in Michigan, USA. A piece of zirconium broke free, blocking the coolant and causing the melting of two assemblies and damage to several others [1]. Detecting blockages from measured temperatures is challenging due to several reasons. First, some rods are used to control the nuclear fission reaction, which can result in lower temperature emissions. Second, the temperature may change according to the power produced, which can be adjusted by lifting or pushing control rods. Third, the model must be accurate to describe regular nuclear core rods temperatures while rejecting abnormal temperature contributions due to blockages.

Let us denote the temperature measured from a rod at location (x, y) as $T_{i;x,y}$. The overall surface of temperature formed by all p rods at instant i is denoted $\mathbf{T}_i \in \mathbb{R}^p$. The goal is to find a model of this surface, which is complex. The accuracy of the model is crucial to distinguish regular changes from changes induced by blockages. The numerical data used in this paper have been measured at a SFR station, though the method can be extended to surveillance of any type of reactor. The specific devices used to measure temperatures were K-type thermocouples, whose transfer function has the form

$$h(t) = \frac{1}{\tau} \exp\left(\frac{-t}{\tau}\right) \quad (1)$$

Since it is not possible to create a real TIB to measure its impact on rods temperatures, numerical simulations were used. The model of the under-construction ASTRID reactor was used [18, 21], taking into account relevant thermodynamic parameters. The result of the simulation does not exactly match the observed temperature due to sampling frequency and thermocouple response. To transform the temperature from the simulation to realistic observable temperature, the K-type thermocouple device response function was applied, and the data was subsampled by averaging temperatures over one second. This gives a model for the anomalous temperature due to the TIB, from the simulated temperature denoted $a(t)$, as follows:

$$\begin{aligned} \mathbf{a} &= (a_1, \dots, a_i), \\ a_i &= \int_{(i-1)\Delta}^{i\Delta} a(t) * h(t) dt \end{aligned} \quad (2)$$

with $*$ the convolution product, Δ the sampling period, and $h(t)$ the thermocouple impulse response provided in Eq.(1).

2.2 Detection of a Known Anomaly

The temperature of each rod is assumed to be statistically independent, and the temperatures of the rods are modelled with a Gaussian distribution. The expected temperature of each rod is denoted $\mu_{i;x,y}$, and the overall temperature of all rods at time index i is denoted \mathbf{T}_i . Under normal operating conditions, the temperature of all rods can be modelled as:

$$\mathbf{T}_i \sim \mathcal{N}\left(\boldsymbol{\mu}_i, \sigma^2 \mathbf{I}_p\right), \quad (3)$$

where $\boldsymbol{\mu}_i$ is the expected value of all rods temperature at index time i , σ^2 is the variance of all rods, and \mathbf{I}_p is the identity matrix.

A common approach is to design a linear model for the temperature of all rods at each time index. This model can be described

as:

$$\mathbf{T}_i \sim \mathcal{N}\left(\mathbf{H}\boldsymbol{\theta}_i, \sigma^2 \mathbf{I}_p\right), \quad (4)$$

where \mathbf{H} is a known full-column rank matrix and $\boldsymbol{\theta}$ is a vector of nuisance parameters.

The proposed methodology relies on invariance theory for hypothesis testing, see [8][Chap. 4]. To apply this theory, we define the column space spanned by \mathbf{H} as $R(\mathbf{H})$. We then project the vector of measured temperature \mathbf{T}_i onto the orthogonal complement $R(\mathbf{H})^\perp$ of the column space $R(\mathbf{H})$. This decomposition allows us to impose restrictions on the expected temperature and distinguish the regular temperature from the temperature due to a TIB. To obtain this decomposition and reject the non-anomalous temperature, we project each observation \mathbf{T}_i onto the $p-q$ dimensional subspace $R(\mathbf{H})^\perp$ as follows:

$$\mathbf{W}\mathbf{T}_i$$

where \mathbf{W} is a matrix composed of the eigenvectors of the projection matrix

$$\mathbf{P}_\mathbf{H}^\perp = \mathbf{I}_p - \mathbf{H}(\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top.$$

The matrix \mathbf{W} verifies the following properties:

$$\mathbf{W}\mathbf{H} = \mathbf{0}, \quad \mathbf{W}^\top \mathbf{W} = \mathbf{P}_\mathbf{H}^\perp, \quad \mathbf{W}\mathbf{W}^\top = \mathbf{I}_{p-q}. \quad (5)$$

Because the Temperature \mathbf{T}_i (3) follows a Gaussian distribution with independent measurements, it is straightforward that the so-called "residuals" follow the following distribution:

$$\mathbf{W}\mathbf{T}_i \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}_{p-q}\right), \quad (6)$$

when no anomaly is present.

Now when an issue occurs, it can be modelled as an additional component in the measurements of temperature:

$$\mathbf{T}_i \sim \mathcal{N}\left(\mathbf{H}\boldsymbol{\theta}_i + \mathbf{a}, \sigma^2 \mathbf{I}_p\right), \quad (7)$$

in which case the distribution of the residuals becomes:

$$\mathbf{W}\mathbf{T}_i \sim \mathcal{N}\left(\mathbf{W}\mathbf{a}, \sigma^2 \mathbf{I}_{p-q}\right), \quad (8)$$

Therefore, the statistical problem of blockage detection can be simplified, after rejection of the non-anomalous background, to the discrimination between the following statistical hypotheses:

$$\begin{cases} \text{Under } \mathcal{H}_0 : \mathbf{W}\mathbf{T}_i \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}_{p-q}\right), \\ \text{Under } \mathcal{H}_1 : \mathbf{W}\mathbf{T}_i \sim \mathcal{N}\left(\mathbf{W}\mathbf{a}, \sigma^2 \mathbf{I}_{p-q}\right), \end{cases} \quad (9)$$

where $\mathbf{W}\mathbf{a}$ is known.

The statistical detection problem (9) corresponds to a test between simple hypotheses in which case it is well known from the Neyman-Pearson Lemma [8, Theorem 3.2.1] that the optimal, most powerful test is the likelihood ratio test. In the case of detection of a known signal $\mathbf{W}\mathbf{a}$ in Gaussian noise, it is straightforward that, for an observation \mathbf{T}_i , the log-likelihood ratio between the alternative hypotheses (9) can be written as:

$$\Lambda(\mathbf{T}_i) = \log \frac{(2\sqrt{\pi}\sigma)^{-(p-q)} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{W}\mathbf{T}_i - \mathbf{W}\mathbf{a}\|^2\right)}{(2\sqrt{\pi}\sigma)^{-(p-q)} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{W}\mathbf{T}_i\|^2\right)} \quad (10)$$

$$= -\|\mathbf{W}\mathbf{T}_i - \mathbf{W}\mathbf{a}\|^2 + \|\mathbf{W}\mathbf{T}_i\|^2 \quad (11)$$

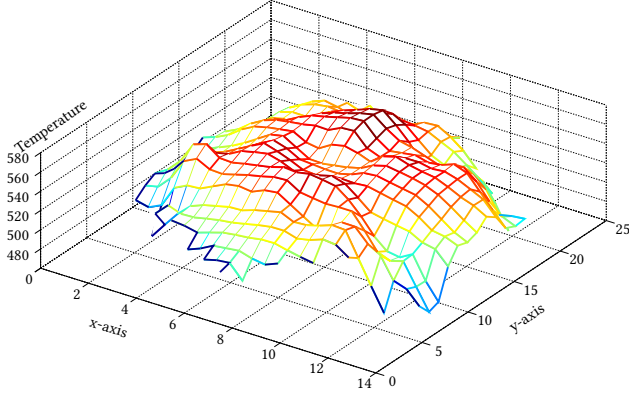


Figure 2: An example of temperature from rods, which illustrates the difficulty of modelling this surface.

with $\|\mathbf{a}\|_2^2 = \mathbf{a}^\top \mathbf{a}$ the L_2 , or Euclidean norm.

It is obvious that the Equation (11) depends on the observation \mathbf{T}_i only through the simplified term:

$$\Lambda^*(\mathbf{T}_i) = (\mathbf{W}\mathbf{T}_i)^\top \mathbf{W}\mathbf{a} = \mathbf{T}_i^\top \mathbf{W}^\top \mathbf{W}\mathbf{a} = \mathbf{T}_i^\top \mathbf{P}_{\mathbf{H}}^\perp \mathbf{a}, \quad (12)$$

which correspond to the well-known matched subspace detectors [16] after rejection of the non-anomalous background via \mathbf{W} .

3 PROPOSED AUTO-ADAPTIVE MODEL OF OBSERVATIONS

In this paper, we propose a polynomial model of degree ℓ for the expectation μ_i of temperature \mathbf{T}_i , which can be written as:

$$\mu_{i;x,y} = \sum_{k=0}^{\ell} \sum_{n=0}^{\ell} d_{k,n} x^k y^n, \quad \forall (x,y) \in \mathcal{X}. \quad (13)$$

This model can be reformulated in matrix form as:

$$\mu_i = \mathbf{G}\mathbf{d}_i, \quad (14)$$

where the columns of the matrix \mathbf{G} span the basis of each polynomial component:

$$\mathbf{G} = \left(\mathbf{1}, \mathbf{x}, \mathbf{y}, \mathbf{x} \cdot \mathbf{y}, \mathbf{x}^2, \dots, \mathbf{x}^\ell \cdot \mathbf{y}^\ell \right). \quad (15)$$

This polynomial model has been shown to be efficient in several signal processing and image processing applications [4, 12, 17, 22, 23]. However, for the present context, a two-dimensional polynomial at every instant is not sufficient to accurately model the non-anomalous temperature. Using a high-degree polynomial would reduce the number of "degrees of freedom" $p-q$ and represent a substantial part of the anomaly due to TIB as regular temperature changes, thus reducing the detection efficiency.

To address this issue, we propose a different approach that takes into account the history of observed temperatures. We exploited the shape of the previous observations with the underlying idea that regular changes in temperature should be smooth and share similar "shapes". To extract the main components of the polynomial modelling errors, we use a simple Principal Component Analysis (PCA) over the polynomial modelling error $\mathbf{W}\mathbf{T}_i$ to extract the c principal components.

The proposed original model is obtained as follows. At each time index i , the fixed polynomial part is used to estimate the coefficients \mathbf{d}_i using the maximum likelihood method, which in the presence of Gaussian i.i.d noise is equivalent to the least square estimation:

$$\tilde{\mathbf{d}}_i = (\mathbf{G}^\top \mathbf{G})^{-1} \mathbf{G}^\top \mathbf{T}_i. \quad (16)$$

The ensuing modeling error residuals are given by:

$$\mathbf{T}_i - \mathbf{G}\tilde{\mathbf{d}}_i = (\mathbf{I}_p - \mathbf{G}(\mathbf{G}^\top \mathbf{G})^{-1} \mathbf{G}^\top) \mathbf{T}_i. \quad (17)$$

Over which the computation of principal components is carried out. Denoting \mathbf{F}_i the matrix of the c principal components (into columns) and observing that the principal components are orthonormal, the maximum likelihood estimation of principal components coefficients is given by:

$$\tilde{\mathbf{u}}_i = \mathbf{F}_i^\top \mathbf{T}_i \quad (18)$$

The final estimation of the regular temperature is obtained as:

$$\tilde{\mathbf{T}}_i = \mathbf{H}_i \tilde{\boldsymbol{\theta}}_i, \quad (19)$$

where the matrix \mathbf{H}_i and the coefficient $\boldsymbol{\theta}_i$ are obtained, respectively, by concatenation column-wise of each part of the model:

$$\mathbf{H}_i = (\mathbf{G}, \mathbf{F}_i) \quad \text{and} \quad \tilde{\boldsymbol{\theta}}_i = \left(\tilde{\mathbf{d}}_i, \tilde{\mathbf{u}}_i \right) \quad (20)$$

4 PROPOSED DETECTION METHOD

Let us recall that, as explained in the previous section 2–3, we propose using a linear model to represent the temperature of rods \mathbf{T}_i :

$$\mathbf{T}_i \sim \mathcal{N} \left(\mathbf{H}_i \boldsymbol{\theta}_i, \sigma^2 \mathbf{I}_p \right).$$

In the present paper, we propose an autoadaptive model to build the linear model \mathbf{H}_i by merging a fixed polynomial part along with the principal components estimated from the previous observations (20).

In addition, we have explained how to deal with the non-anomalous "background" $\mathbf{H}_i \boldsymbol{\theta}_i$ by projecting the measurements onto \mathbf{W}_i which spans the orthogonal complement of the matrix \mathbf{H}_i hence \mathbf{W}_i consists of the eigenvectors of

$$\mathbf{P}_{\mathbf{H}_i}^\perp = \mathbf{I}_p - \mathbf{H}_i (\mathbf{H}_i^\top \mathbf{H}_i)^{-1} \mathbf{H}_i^\top.$$

Within the practical context of critical infrastructure and monitoring, a maximal detection delay of M temperature measurements is prescribed. Hence, at each instant i it is proposed to inspect the M last samples. Additionally, the N previous samples $\{\mathbf{T}_{i-M-N+1}, \dots, \mathbf{T}_{i-M}\}$ are used to build the proposed adaptive linear parametric model \mathbf{H}_i and to estimate its parameters, as described in Section 3.

Besides, thanks to the numerical simulation described in Section 2.1, the model of temperature drift, or anomaly, after the occurrence of a TIB \mathbf{a}_i is known with the highest accuracy. Therefore, both \mathbf{W}_i and $\mathbf{a}_{i-\nu}$ are known.

In other words, the proposed method relies thus on two windows of fixed length, the first being used to estimate the principal components of the temperature and the later being used for the detection, see Figure 3.

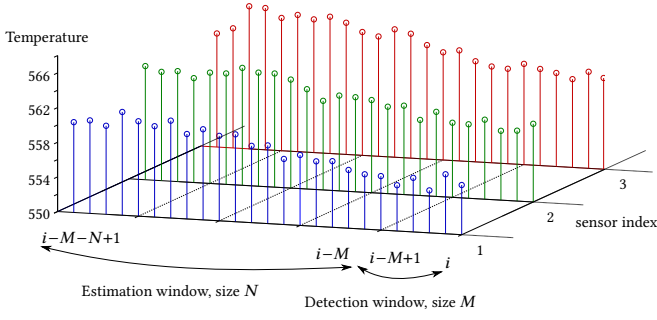


Figure 3: Illustration of the proposed method based on two non-overlapping windows; the first is used to allow adaptivity of the linear model, while the second is used for TIB detection.

4.1 Statistical Properties of the Proposed Method

For clarity and simplicity, let us define, for all $j \in [i - M + 1, \dots, i]$, the normalized sufficient statistics Λ_i as follows

$$\Lambda_i^* = \Lambda^*(T_i) = \frac{1}{\sigma} T_i^\top P_{H_i}^\perp \mathbf{a}. \quad (21)$$

It is straightforward from relations (3), (4), (9) and (21) to establish the distribution of Λ_i when the blockage is not present:

$$\text{Under } \mathcal{H}_0 : \Lambda_i^* \sim \mathcal{N}\left(0, \|\mathbf{a}_i^\perp\|_2^2\right) \quad (22)$$

with

$$\|\mathbf{a}_i^\perp\|_2^2 = \|\mathbf{W}_i \mathbf{a}\|_2^2 = \mathbf{a}^\top P_{H_i}^\perp \mathbf{a}. \quad (23)$$

On the opposite, in the presence of a blockage, it immediately follows from Equations (3), (4), (9) and (21) that

$$\text{Under } \mathcal{H}_1 : \Lambda_i^* \sim \mathcal{N}\left(\sigma^{-1} \|\mathbf{a}_i^\perp\|_2^2, \|\mathbf{a}_i^\perp\|_2^2\right). \quad (24)$$

From the previous relations (21)–(24) and assuming that the linear model presented in section 3 holds true, it is straightforward to establish the statistical properties of the likelihood ratio test:

$$\delta_i = \begin{cases} \mathcal{H}_0 : & \text{if } \Lambda_i^* \leq \tau \\ \mathcal{H}_1 : & \text{if } \Lambda_i^* > \tau \end{cases} \quad (25)$$

where τ is the decision threshold.

Indeed, under hypothesis \mathcal{H}_0 it immediately follows that the following decision threshold:

$$\tau_0 = \|\mathbf{a}_i^\perp\|_2 \Phi^{-1}(1 - \alpha_0) \quad (26)$$

where Φ and Φ^{-1} represent the cumulative distribution function of the normal distribution and its inverse, allows complying with a prescribed false alarm rate α_0 .

Similarly, it immediately follows from the distribution of the test statistics (24) that the power function β , sometimes also referred to as the sensitivity, is given by:

$$\beta = 1 - \Phi\left(\frac{\tau_0 - \sigma^{-1} \|\mathbf{a}_i^\perp\|_2^2}{\|\mathbf{a}_i^\perp\|_2}\right) \quad (27)$$

$$= 1 - \Phi\left(\Phi^{-1}(1 - \alpha_0) - \sigma^{-1} \|\mathbf{a}_i^\perp\|_2\right). \quad (28)$$

We shall see in the next Section 5 that the proposed model is accurate enough to verify in practice that the statistical performance as established in (26)–(28) holds true with real-world data.

4.2 Proposed Sequential Detection Methodology

Another originality of the present paper is that it is not aimed at detecting the blockage from a single measurement but rather to detect the occurrence of this event in less than a few seconds. In other words, over a set of successive temperature measures. In the literature this problem of sequential analysis is referred to as a change-point detection see, for instance, [20] and the reference therein for a complete introduction. To this end we propose using the decision statistics of the proposed LR test (21) as an input for the well-known CUSUM (cumulative sum) sequential method initially developed by Page [14]. Let us denote $\Lambda_1^*, \dots, \Lambda_I^*$ the consecutive values of the likelihood ratio. The idea of the CUSUM is to compute a cumulative sum of the decision statistics Λ_i^* . After the occurrence of the change point, that is after the blockage occurs in our case, the average value of the LR statistics shall increase. Therefore the CUSUM is computed as:

$$S_i = \max(0; S_i + \Lambda_i^* - c), \quad (29)$$

$$S_0 = 0. \quad (30)$$

In the relation (30) one can note that the constant value c is subtracted from the LR value. This constant represents the minimum value above which it is considered that a blockage has been detected. In addition one can also note in the equation (30) that the CUSUM value S_i is reset to 0 when it is negative. This is a key concept of the CUSUM, which allows detecting the change point with a minimal delay because samples without blockage do not reduce the CUSUM value, which is assumed to remain at 0 before the occurrence of the event we want to detect, in our case, the total and instantaneous blockage.

On the one hand, it has been shown that the CUSUM minimizes the average detection delay (??) for a fixed average false-alarm time [9, 11]. On the other hand, it would be interesting to minimize the maximal detection, instead of the average, in order to guarantee a maximal delay of detection. However, to the best of our knowledge, there isn't any work that proposes such a sequential detection system.

5 NUMERICAL EXPERIMENTATION AND RESULTS

The real data used in the present work were provided by the French Commission of Atomic and Alternative Energies (CEA) and obtained from the Sodium-cooled Fast Reactor (SFR) of the experimental French Phenix Nuclear Reactor. Two datasets of one week have been provided, one recorded from February 16th 2009 to February 22nd 2009 and the second one starting on March 02nd 2009 ; those give us a total of two-time 201 600 data for each of the 120 rods. Though the original sampling rate was three seconds, we considered that the sampling frequency was 1Hz because, first, it corresponds to the sampling frequency of the newest nuclear power plants, hence its relevance for the design of the next generation ASTRID reactor.

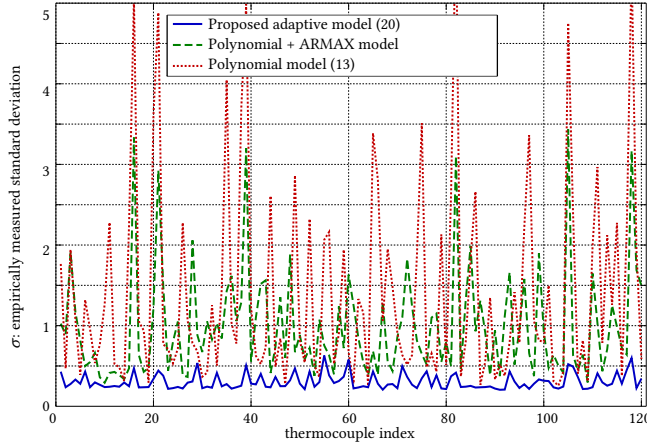


Figure 4: Standard deviation of residuals for each thermocouple. The smallest standard deviation of the proposed adaptive model shows its accuracy.

First of all, let us start by verifying the accuracy and interest of the proposed adaptive model; we compare it with its “fixed” part and with the ARMAX spatial and temporal model proposed in [13]. Note that this model is, to the best of our knowledge, the only comparable method proposed in the literature. To give briefly a description of the method proposed in [13], let us recall that the ARMAX can be described with the following relations:

$$\begin{cases} \mathbf{T}_i &= \mathbf{G}\theta_i + \mathbf{R}_i \\ (1 - B_{x,y}\kappa_{x,y})R_{i;x,y} &= (1 - B_{x,y}\eta_{x,y})\xi_{i;x,y} \end{cases} \quad (31)$$

where $\xi_{i;x,y} \sim \mathcal{N}(0, \sigma^2)$ is a Gaussian random variable that represents the noise contribution on temperature measurement for rod at position (x, y) .

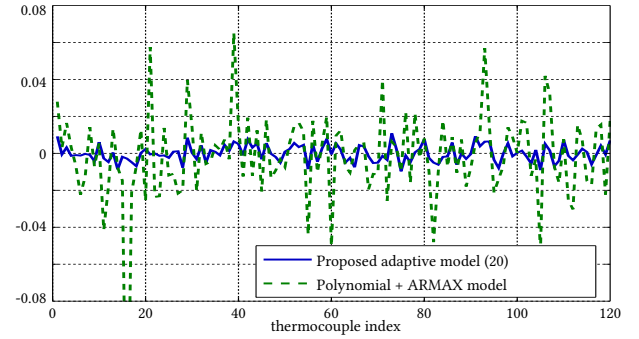
We have implemented the method proposed in [6] for online estimation of ARMAX parameters with forgetting factors. We set the forgetting factor to $0.999 = 1 - 0.001$, which has been found empirically to give the best overall results.

The first results compare the statistical properties of temperature estimation. Figure 4 compares the variance of this estimation through the standard deviation σ of all the residuals \mathbf{T}_i for each rod and for three models: the proposed adaptive model, the fixed polynomial part of this model, and the ARMAX model. A small variance of residuals represents a high accuracy of the model. Figure 4 shows that the proposed adaptive model offers a much smaller variance of residuals than the fixed polynomial part of this model, even when using an ARMAX model.

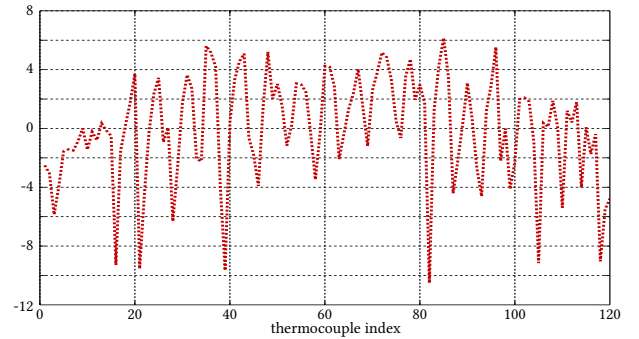
This result emphasizes the interest in adding the adaptive part F_i that allows, with a rather sparse model of 13 basis vectors for 120 rods, to get an estimation of rods temperatures with smaller variance compared to the alternative approaches.

Interestingly, one can note that for the proposed adaptive model, the mean standard over all residuals from all rods is about 0.2625°C , which is in line with reference noise values [3, Chap.5]. Besides, one can also note that several rods have a larger variance than the others, which is due, at least in part, to the provided dataset.

Similarly, figures 5a and 5b present the other important properties of temperature estimation, namely the empirical bias. Figure 5a



(a) Bias of estimated temperature by rods for proposed adaptive model.



(b) Bias of estimated temperature by rods using only the fixed polynomial part of the model.

Figure 5: Comparison of empirical bias on estimated temperature using the proposed adaptive model and the fixed polynomial part of this model.

shows the bias for the proposed adaptive model and the fixed polynomial part with ARMAX, while Figure 5b presents only the bias for the fixed polynomial model alone. The choice of using two figures for those results has been driven by the difference in the scale of the bias.

The results emphasize the poor accuracy of the polynomial model to describe the temperature “surface”, which is very complex. The least square estimation of the polynomial that best fits the temperature surface makes, in mean, large errors for most rods, which are shown by a large bias. Those results also confirm again the interest in the proposed adaptive model that has obviously the smallest bias with negligible values. Together with the previous figure, this shows the quality of temperature estimation obtained using the proposed model that allows getting an estimator with both small bias and small variance.

Of course, one of the main interesting results is the comparison between the theoretically established performance (26) and (??). First of all, let us present the criteria which the most important in the operational context, namely the empirically false-alarm probability and its comparison with the theoretical values (26). Indeed, being able to master the false-alarm rates with high accuracy over numerical simulations with real-world data constitutes an important challenge. The Figure 7 shows the false alarm rate as a function of the decision threshold τ used in the statistical test (25). On this

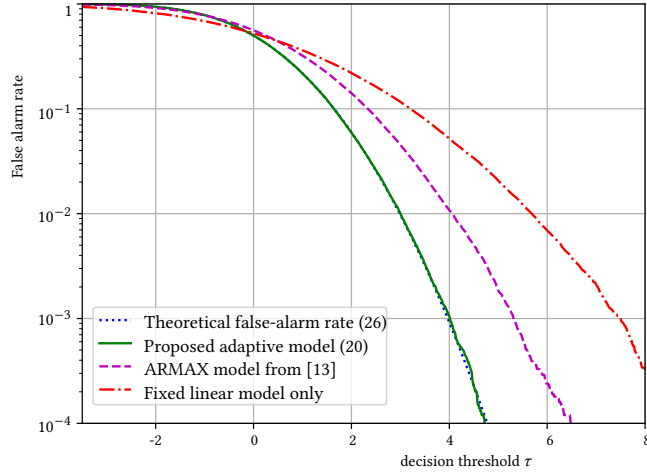


Figure 6: Empirical and theoretical false alarm rate as a function of the decision threshold τ . The comparison between the theoretical results (26) and the empirical results obtained with the proposed model shows the accuracy of the rejection of non-anomalous background.

figure, we also compare the results obtained with the “fixed” linear part of our model, in order to show the relevance of the autoadaptive methods based on PCA as well as the ARMAX model proposed in [13]. One can note that for a high decision threshold, the LR test based on the proposed original autoadaptive model always has a much lower false positive rate. More importantly, we also plot the theoretical false-positive rate as a function of the decision threshold as established in (26). Note that, even though we used only 500,000 simulations based on real data, the false positive rate obtained with the proposed model remains very close to the theoretical one even for false positive rates as low as 10^{-4} . In the operational context, this is very important, as it demonstrates the ability to control the occurrence of false alarms, which poses important practical problems in real-world conditions.

In addition to the Figure 6, the Figure 7 presents the statistical performance of the test as a ROC (Receiver Operating Characteristics) curve, which shows the power function, which is the true-positive rate also referred to as the sensitivity of the test, as a function of the false alarm or false positive rate. Once again, in the Figure 7 the performance obtained with the proposed original autoadaptive model is contrasted with the results obtained with the fixed linear part of our model as well as with the ARMAX model proposed in [13]. Even though we have to simulate the blockage and reducing its effect, using a multiplicative factor $1/6$, it is obviously the LR test, based on the proposed model, performs much better than the two other alternative approaches. Again, compare the empirical power function with the theoretically established one (28). Interestingly, empirical power function remains very close to the theoretical one even for false alarm, or false positive, rates as low as 10^{-4} which represent only 50 observations in our simulation with 500,000 data from the real temperature measurements.

These results show the relevance and the accuracy of the proposed autoadaptive model as well as the sharpness of the theoretical findings.

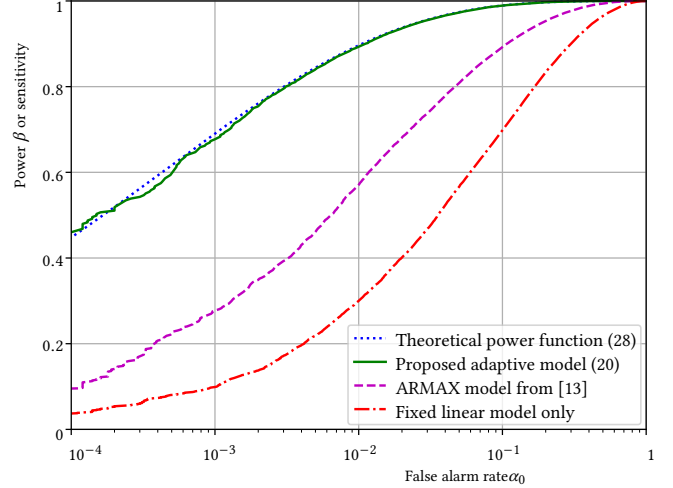


Figure 7: Receiver Operating Characteristic (ROC) curves comparing the performance of blockage with different statistical models. The comparison with the theoretical power (28) shows the accuracy of the proposed autoadaptive model.

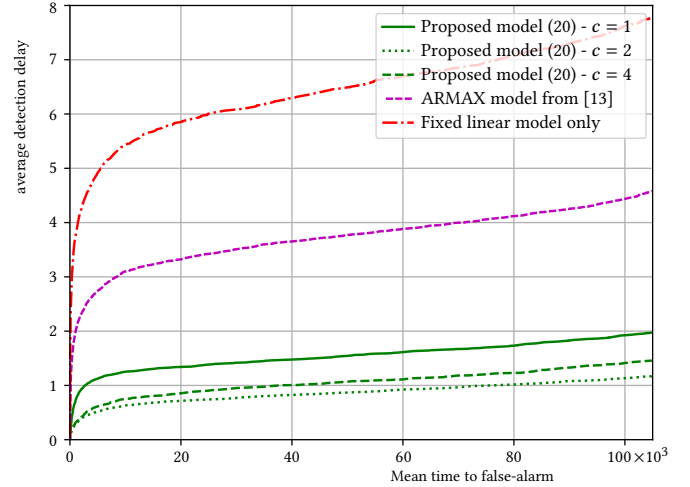


Figure 8: Average detection delay (in samples or seconds) as a function of the mean time to false alarm.

The last results we would like to present concern about the detection delay and the mean time to false alarm of the proposed sequential detection scheme based on the CUSUM presented in Section 4.2. To this end, Figure 8 shows the comparison of the detection delay for the proposed autoadaptive model with the ARMAX model proposed in [13] and the fixed (non-adaptive) polynomial part of the proposed model. One course if one wishes to increase the average time before a false alarm occurs, it shall increase the detection threshold of the CUSUM (30) statistics and hence must accept a higher delay of detection after a blockage occurs. It is obvious in Figure 8 that the proposed autoadaptive linear model allows achieving a much smaller detection delay of the blockage for the same average time a before a false alarm occurs. For the sake of comparison, for the proposed adaptive model, the Figure 8 shows the results obtained with different constant c of the CUSUM statistics, see Eq. (30). The best results are obtained with $c = 2$ for

the proposed autoadaptive model. Note that, on the opposite for the two other models used for the comparison, we used the constant c that gives best empirical results, which was around $c = 3$.

6 CONCLUSIONS AND FUTURE WORKS

This paper addresses the problem of a known anomaly in a complex noisy environment. The approach proposed in the present works consists in rejecting non-anomalous observations using a linear parametric model. The main originality of our work is that, in order to deal with a complex and changing environment, it is proposed to include in the linear parametric model principal components from the previous observations in order to increase the flexibility of the model and to allow representing the ever-changing complex environment with high accuracy. The proposed autoadaptive model is used in a statistical test for the application of the detection of an anomaly in the temperature of a nuclear power plant because of the blockage of the coolant. Numerical results show the relevance of the proposed model and its very high accuracy. Indeed, the proposed autoadaptive model allows establishing analytically the statistical performance of the ensuing LR test in terms of false positive and true positive probability. A wide range of numerical results based on real-world data from an experimental power plant show the relevance of the proposed methodology and the sharpness of the theoretical findings.

Possible future works include the comparison with deep learning-based methods and the design of a sequential detection method whose statistical performance can be analytically established.

REFERENCES

- [1] 2011. *Experimental Facilities for Sodium Fast Reactor Safety Studies*. Technical Report. Task Group on Advanced Reactors Experimental Facilities (TAREF), Committee on the Safety of Nuclear Installations (CSNI), Nuclear Energy Agency (NEA). 144 pages.
- [2] S Beils, B Carlucci, N Devictor, GL Fiorini, and JF Sauvage. 2009. Safety for the future Sodium cooled Fast Reactors. In *Proc. of International Conference on Fast Reactors and Related Fuel Cycles FR09, Kyoto, Japan*.
- [3] Peter RN Childs. 2001. *Practical temperature measurement*. Butterworth-Heinemann. <https://doi.org/10.1016/B978-075065080-9/50005-8>
- [4] Rémi Coganne, Guillaume Doyen, Nisrine Ghabban, and Badis Hammi. 2018. Detecting Botclouds at Large Scale: A Decentralized and Robust Detection Method for Multi-Tenant Virtualized Environments. *IEEE Transactions on Network and Service Management* 15, 1 (2018), 68–82. <https://doi.org/10.1109/TNSM.2017.2785628>
- [5] Vincent Gatin. 2010. Études d'accidents du Réacteur à Neutrons Rapides à caloporteur Gaz (RNR-G) en support aux études de sûreté de son combustible.
- [6] Satoru Goto, Masatoshi Nakamura, and Katsuji Uosaki. 1995. On-line spectral estimation of nonstationary time series based on AR model parameter estimation and order selection with a forgetting factor. *IEEE Transactions on Signal Processing* 43, 6 (1995), 1519–1522.
- [7] Laurent Hartert, Danielle Nuzillard, and Jean-Philippe Jeannot. 2013. Dynamic detection of nuclear reactor core incident. *Signal Processing* 93, 2 (2013), 468 – 475. <https://doi.org/10.1016/j.sigpro.2012.08.021>
- [8] E.L. Lehmann and J.P. Romano. 2005. *Testing Statistical Hypotheses, Second Edition* (3rd ed.). Springer.
- [9] Gary Lorden. 1971. Procedures for reacting to a change in distribution. *The Annals of Mathematical Statistics* (1971), 1897–1908.
- [10] Sinuhe Martinez-Martinez, Nadhir Messai, Jean-Philippe Jeannot, and Danielle Nuzillard. 2015. Two neural network based strategies for the detection of a total instantaneous blockage of a sodium-cooled fast reactor. *Reliability Engineering & System Safety* 137 (2015), 50 – 57. <https://doi.org/10.1016/j.res.2014.12.003>
- [11] George V Moustakides. 1986. Optimal stopping times for detecting changes in distributions. *the Annals of Statistics* 14, 4 (1986), 1379–1387.
- [12] Tan Nguyen, Remi Coganne, and Guillaume Doyen. 2015. An optimal statistical test for robust detection against interest flooding attacks in CCN. In *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. 252–260. <https://doi.org/10.1109/INM.2015.7140299>
- [13] Igor Nikiforov, Fouzi Harrou, Rémi Coganne, Pierre Beausery, Edith Grall, Blaise Kevin Guépié, Lionel Fillatre, and Jean-Philippe Jeannot. 2020. Sequential detection of a total instantaneous blockage occurred in a single subassembly of a sodium-cooled fast reactor. *Nuclear Engineering and Design* 366 (2020), 110733. <https://doi.org/10.1016/j.nucengdes.2020.110733>
- [14] ES Page. 1954. Continuous inspection schemes. *Biometrika* (1954), 100–115.
- [15] K. Paumel, J.P. Jeannot, T. Jeanne, G. Laffont, M. Vanderhaegen, and N. Mascret. 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 *Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA), 2013 3rd International Conference on*. 1–12. <https://doi.org/10.1109/ANIMMA.2013.6727985>
- [16] Louis L Scharf and Benjamin Friedlander. 1994. Matched subspace detectors. *IEEE Transactions on signal processing* 42, 8 (1994), 2146–2157.
- [17] Vahid Sedighi, Rémi Coganne, and Jessica Fridrich. 2015. Content-Adaptive Steganography by Minimizing Statistical Detectability. *Information Forensics and Security, IEEE Transactions on* (in press) (2015).
- [18] Nathalie Seiler, Jean-Baptiste Droin, Frederic Bertrand, and Rémi Clavier. 2023. Design-oriented simulation tools in french SFR projects. In *Technical Meeting on the Safety Approach for Liquid Metal Cooled Fast Reactors and the Analysis and Modelling of Severe Accidents*. IAEA-CN.
- [19] Alexander Stanculescu. 2018. GIF R&D outlook for generation IV nuclear energy systems: 2018 update. In *Proceedings of the Generation IV International Forum, Paris, France*. 16–18.
- [20] A. Tartakovsky, I. Nikiforov, and M. Basseville. 2014. *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. Taylor & Francis.
- [21] Stéphanie Tillement and Frédéric Garcias. 2021. ASTRID, back to the future: bridging scales in the development of nuclear infrastructures. *Nuclear Technology* 207, 9 (2021), 1291–1311.
- [22] Karim Tout, Rémi Coganne, and Florent Retrait. 2018. Statistical decision methods in the presence of linear nuisance parameters and despite imaging system heteroscedastic noise: Application to wheel surface inspection. *Signal Processing* 144 (2018), 430–443. <https://doi.org/10.1016/j.sigpro.2017.10.030>
- [23] Hao Yin, Chuang Lin, Berton Sebastien, Bo Li, and Geyong Min. 2005. Network traffic prediction based on a new time series model. *International Journal of Communication Systems* 18, 8 (2005), 711–729.

submitted 20 Decembre 2024