Spatio-Temporal Decision Fusion for Quickest Fault Detection Within Industrial Plants: The Oil and Gas Scenario

Gianluca Tabella*, Domenico Ciuonzo†, Nicola Paltrinieri‡, Pierluigi Salvo Rossi*§
*Dept. Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway
†Dept. Electrical Engineering and Information Technologies (DIETI), University of Naples "Federico II", Naples, Italy
‡Dept. Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway
§Dept. Gas Technology, SINTEF Energy Research, Norway
Email: gianluca.tabella@ntnu.no; domenico.ciuonzo@unina.it; nicola.paltrinieri@ntnu.no; salvorossi@ieee.org

***Abstract*—In this work, we present a spatio-temporal decision fusion approach aimed at performing quickest detection of faults within an Oil and Gas subsea production system. Specifically, a sensor network collectively monitors the state of different pieces of equipment and reports the collected decisions to a fusion center. Therein, a spatial aggregation is performed and a global decision is taken. Such decisions are then aggregated in time by a post-processing center, which performs quickest detection of system fault according to a Bayesian criterion which exploits change-time statistical distributions originated by system components' datasheets. The performance of our approach is analyzed in terms of both detection- and reliability-focused metrics, with a focus on (fast & inspection-cost-limited) leak detection in a real-world oil platform located in the Barents Sea.**

***Index Terms*—data fusion, distributed detection, maintenance, monitoring, reliability, wireless sensor network.**

I. Introduction

THE last decade has seen the growth of Wireless Sensor Networks (WSNs) and their use in monitoring applications. In particular, the task of event detection and localization has received large attention, especially in relation to the design of barriers for safety-critical systems. WSNs are usually made of low-cost devices having the task of monitoring the surrounding environment. In order to lower communication and processing costs, the sensors are often designed to transmit binary decisions to a Fusion Center (FC), which has the task to collect the local decisions and formulate a global decision on whether the event of interest is occurring. When an adverse event is detected, the FC produces an alarm so that proper actions can be taken to mitigate the event's consequences.

This scenario particularly applies to the process, energy, and manufacturing industry, where the failure of a piece of equipment could compromise the safety of the workers and the environment. Indeed, this may result in high costs as well as missed revenues due to unplanned shutdowns [1].

This research is a part of BRU21 – NTNU Research and Innovation Program on Digital and Automation Solutions for the Oil and Gas Industry (www.ntnu.edu/bru21).

Event detection via WSNs for industrial applications has been an object of study and many architectures have been analyzed and proposed, lately with a focus on underwater environments [2], [3]. One of the vital issues of the industry is the detection of equipment failures that can lead to dangerous losses of containment, especially in those environments where inspections are highly costly.

Various algorithms for detection of non-cooperative targets via distributed WSNs are presented in [4], where the Generalized Likelihood Ratio Test (GLRT), the Bayesian approach, and a hybrid GLRT/Bayesian detector are used to approach the problem. In [5] the sub-optimal *Counting Rule* (CR) is extended with an ordering-based where highly-informative sensors have a higher priority in the transmission to the FC, showing the same performances as the classical CR with fewer transmissions. In [6], the CR is applied to the case of radiation detection, while in [7], [8] the same rule is employed for subsea oil spill detection. In [9], [10] the work is extended including a modified version of the *Chair-Varshney Rule* incorporating in the design of the WSN the positions and the failure rates of those items susceptible to failure.

Nevertheless, no approach has been provided that successfully incorporates information on the reliability of the monitored system in the design of the detection algorithm. Such information is often available and regards critical items whose failure is to be detected exploiting their emitted signal. Data such as positions, failure rates, and failure models constitute prior information that can be embodied in a Bayesian framework.

Accordingly, the *main contributions* of this work are summarized as follows. We propose a spatio-temporal decision fusion approach aimed at performing quickest detection of faults within a critical system. More specifically, a sensor network collectively monitors the state of different pieces of equipment and reports their decisions to a FC. Herein, a *spatial aggregation* is performed and an optimal per-sample decision is performed. Such decisions are then aggregated *in time* by a Post-Processing Center (PPC), which performs quickest detection of faulty system according to a Bayesian criterion and

exploits change-time statistical distributions driven by system components' datasheets. The separation between the FC and PPC allows for system modularity and implementation of the two components via appealing edge-cloud architectures [11]. The results of the proposed approach are analyzed focusing on a real Oil and Gas setup, namely the Goliat FPSO oil production system. Results, both in terms of (i) detection and (ii) reliability-focused metrics, highlight the appeal of the proposed approach and the additional benefit of temporal aggregation (as opposed to the sole spatial aggregation).

The rest of the manuscript is organized as follows. Sec. II describes the system model considered, whereas Sec. III introduces the proposed decision fusion approach for quickest fault detection. Our approach is then numerically validated on a real case study in Sec. IV. Finally, Sec. V ends the paper with some pointers to future directions of research.

Notation – vectors are denoted with bold letters; $(\cdot)^T$, and $\|\cdot\|$ denote transpose and Euclidean norm operators, respectively; $\hat{a}$, $\mathbb{E}\{a\}$, and $\mathbb{E}\{a|b\}$ denote an estimate of the random variable a, its expectation, and its conditional expectation given the random variable b, respectively; $\Pr(\cdot)$ and $p(\cdot)$ denote probability mass functions (pmfs) and probability density functions (pdfs), while $\Pr(\cdot|\cdot)$ and $p(\cdot|\cdot)$ their corresponding conditional counterparts; $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2; $\mathcal{Q}(\cdot)$ is the complementary cumulative distribution function (ccdf) of the standard normal distribution; $\text{Exp}(\lambda)$ denotes an exponential distribution with rate λ; $\text{Gamma}(\alpha, \beta)$ denotes a Gamma distribution with shape α and rate β; the symbol $\sim$ means "distributed as".

II. System Model

A. Failure Model

The monitored system consists of $m = 1, 2, \ldots, M$ points, with mth point located at position $\boldsymbol{h}_m$. These points are *items* (pieces of equipment at locations of interest) and their individual state at time t is described via the following variable:

$$\mathcal{H}_m(t) = \begin{cases} 0, & m\text{th item in } \textbf{active} \text{ state at time } t \\ 1, & m\text{th item in } \textbf{failed} \text{ state at time } t \end{cases}, \quad (1)$$

where *active* means that the item is working the way it was intended and no action should be taken, whereas *failed* means that the item is in a condition requiring maintenance. Moreover, we define the state variable for the *entire system*:

$$\mathcal{H}(t) = 1 - \prod_{m=1}^{M} (1 - \mathcal{H}_m(t)) = \begin{cases} 0, & \text{system in } \textbf{active} \text{ state at time } t \\ 1, & \text{system in } \textbf{failed} \text{ state at time } t \end{cases}, \quad (2)$$

meaning that the system is considered in failed state when at least one of its items is in such state and that the failures are independent. An item in failed state maintains its condition until maintenance is performed. While in failed state, the item generates a signal (any significant change in a measurable variable caused by the failed state). Once the signal is detected, an inspection is performed to assess the condition of the whole system, and maintenance is eventually performed on all items in failed state.

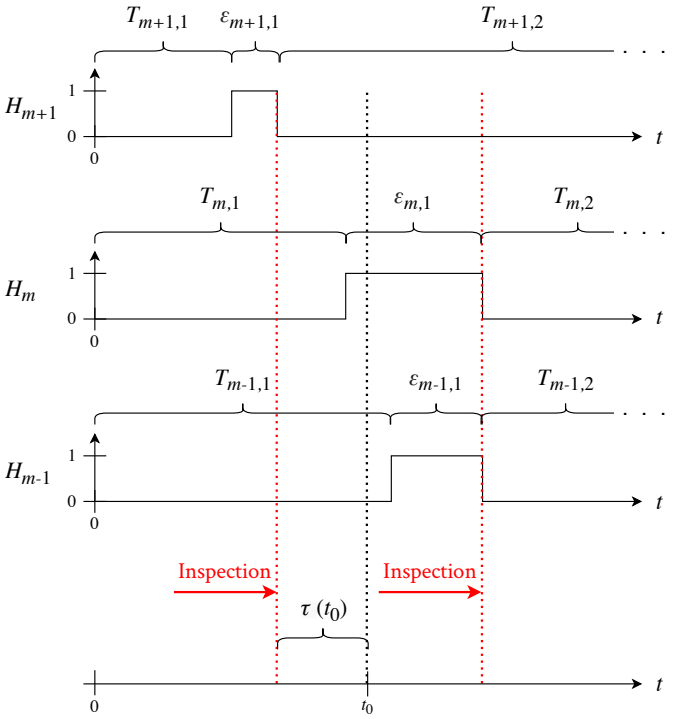

Fig. 1. Failure model (inspection and maintenance time are neglected).

The failure of the mth item is modeled as a *homogeneous Poisson process* with *failure rate* λ_m (see Fig. 1). Let us define $T_{m,j}$ as the time spent by the mth item between the $(j-1)$th and the jth failure, then $T_{m,j} \sim \text{Exp}(\lambda_m)\ \forall j \in \mathbb{N}$. In addition, we define $T^*_{m,j} \triangleq T_{m,j} + \varepsilon_{m,j}$, where $\varepsilon_{m,j}$ is the time spent before the failed state is acknowledged by an inspection. At time t, we can also define $\tau(t)$ as the amount of time since the last inspection. Because of the *memoryless property* of the homogeneous Poisson process, maintenance can be seen either as component repair or substitution. As a consequence of the failure model we can obtain its *failure function*:

$$F_m(t) = \Pr(\mathcal{H}_m(t) = 1) = 1 - e^{-\lambda_m \tau(t)}\,. \quad (3)$$

Next, we can obtain the failure function of the whole system at a given time t:

$$F(t) = \Pr(\mathcal{H}(t) = 1) = 1 - \prod_{m=1}^{M} (1 - F_m(t))\,. \quad (4)$$

Moreover, for sufficiently small λ_m's, we can simplify the model:

$$F(t) \approx \sum_{m=1}^{M} F_m(t)\,. \quad (5)$$

Such approximation implies *disjoint failures*, meaning that at any time t, *at most* one item is in failed state and will be exploited in the design of the detectors.

The exact failure rate of an item λ_m is typically *unknown*, however, estimates (average values and their variances) are usually available for categories of items. Therefore, the proposed method will treat the failure rate of each item as a random variable whose realization must be estimated within a Bayesian framework during the item lifetime.

Finally, as the system monitoring is performed at constant intervals of length Δt (except during inspection and maintenance), we generically consider the processing at nth discrete time, where $n = 1$ corresponds to the first algorithm instance since the last inspection.

B. Signal Model

When an item is in failed state, a signal is generated. Such signal is sensed by K sensors deployed in the environment. The model of the received signal $y_k[n]$ at the kth sensor during the nth discrete time is the following:

$$y_k[n] = \sum_{m=1}^{M} s_{m,k}[n] + w_k[n] \,, \tag{6}$$

where $s_{m,k}[n]$ and $w_k[n] \sim \mathcal{N}(0,\, \sigma^2_{w,k})$ represent the received signal from the mth item and the Additive White Gaussian Noise (AWGN), respectively. $s_{m,k}[n]$ and $w_k[n]$ are assumed statistically independent due to the spatial separation of the sensors. More specifically, $s_{m,k}[n]$ is assumed to have the following shape:

$$s_{m,k}[n] = \begin{cases} 0 & \text{if } \mathcal{H}_m[n] = 0 \\ \xi_m[n]\; g(\boldsymbol{x}_k, \boldsymbol{h}_m) & \text{if } \mathcal{H}_m[n] = 1 \end{cases} , \tag{7}$$

where $\xi_m[n] \sim \mathcal{N}(0,\, \sigma^2_{\xi,m})$ represents the emitted signal by the mth item in failed state at a reference length (ℓ_{ref}). The values of $\sigma^2_{\xi,m}$ and $\sigma^2_{w,k}$ are assumed to be *known*. Finally, $g(\boldsymbol{x}_k, \boldsymbol{h}_m)$ represents the Amplitude Attenuation Function (AAF), depending on the distance between the kth sensor position ($\boldsymbol{x}_k$) and the mth item position ($\boldsymbol{h}_m$).

It is worth noticing that the considered model can be reduced to a binary hypothesis as a consequence of the simplification introduced in Eq. (5) which excludes the possibility of more than one item to be in failed state at a given moment. Accordingly, it holds:

$$\begin{cases} \mathcal{H}[n] = 0: & y_k[n] = w_k[n] \\ \mathcal{H}[n] = 1: & y_k[n] = \xi_m[n]\; g(\boldsymbol{x}_k, \boldsymbol{h}_m) + w_k[n] \end{cases} . \tag{8}$$

Consequently, we can write the statistics of the signal:

$$\begin{cases} y_k|\mathcal{H} = 0 \;\sim\; \mathcal{N}\left(0,\, \sigma^2_{w,k}\right) \\ y_k|\mathcal{H} = 1 \;\sim\; \mathcal{N}\left(0,\, \sigma^2_{\xi,m}\, g^2(\boldsymbol{x}_k, \boldsymbol{h}_m) + \sigma^2_{w,k}\right) \end{cases} . \tag{9}$$

In both Eqs. (8) and (9), only the failure of the generic mth item can be the cause of the system being in failed state.

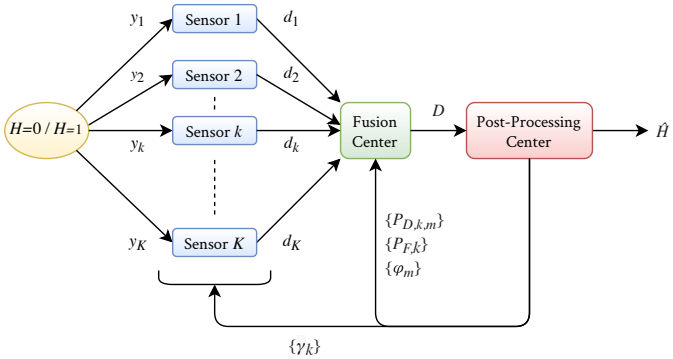

Fig. 2. Distributed Wireless Sensor Network.

C. Wireless Sensor Network Model

The proposed distributed WSN architecture (see Fig. 2) is made of K sensors, one FC, and a PPC. The sensors sense the environment at fixed time intervals of length Δt to detect whether the system is in active ($\mathcal{H}[n] = 0$) or failed ($\mathcal{H}[n] = 1$) state[1]. The kth sensor measures the signal $y_k[n]$ and individually computes a (time-dependent) decision statistic $\Lambda_{k,n}(y_k[n])$, which is then compared to a threshold $\gamma_k[n]$. The sensor takes a local decision $d_k[n] = i \in \{0, 1\}$ if $\mathcal{H}[n] = i$ is declared.

The vector of local decisions $\boldsymbol{d}[n] = \begin{bmatrix} d_1[n] & \cdots & d_K[n] \end{bmatrix}^T$ is collected and processed at the FC for a global decision $\mathcal{D}[n] = i \in \{0, 1\}$ if $\mathcal{H}[n] = i$ is declared. In addition to being spectrally efficient, as only 1-bit communication is required on the reporting channel between the sensor and the FC, such a system is highly energy-efficient when On-Off Keying (OOK) is employed for communicating the local decisions.

Both sensors and the FC perform a *Maximum Likelihood (ML) Detection* on the simplified binary hypothesis of Eq. (8), assuming no prior knowledge on the probability of the events $\mathcal{H}[n] = 0$ and $\mathcal{H}[n] = 1$, and considering every measurement statistically independent.

The PPC collects $\mathcal{D}[1], \ldots, \mathcal{D}[n]$ and integrates the knowledge of the failure model accounting for the signal model of Eq. (6) and (7), and takes a final decision $\widehat{\mathcal{H}}[n]$ through a *Maximum A-Posteriori (MAP) Detection*, with $\widehat{\mathcal{H}}[n] = 1$ triggering inspection operations. The PPC has also the task to continuously communicate to the sensors an updated value of their respective γ_k and a list of parameters necessary for the FC to perform the global detection task. Moreover, at each confirmed failure of the mth item, the PPC will generate a new estimate of λ_m for all $m = 1, \ldots, M$.

III. Proposed Reliability-Based Fusion Approach

This section describes the *three* functional blocks constituting the proposed approach. Specifically, Sec. III-A provides details for the local detection approach at each sensor, while Sec. III-B describes the fusion rule implemented at the FC,

[1]The analysis related to the sampling frequency is not considered in the present work.

implementing spatial processing. Finally, Sec. III-C describes the (a) temporal integration of FC decisions and (b) online item failure rate estimation made by the PPC.

A. Local Detection

At local and sample level, the optimal test is a Likelihood Ratio Test (LRT) associated with the binary hypothesis of Eq. (8) and based on $y_k[n]$, where the position of the failed item is marginalized by using its prior distribution. Specifically, for kth sensor at time n, it holds:

$$\Lambda_{k,n}(y_k[n]) = \frac{p(y_k[n] \,|\, \mathcal{H}[n]=1)}{p(y_k[n] \,|\, \mathcal{H}[n]=0)} \tag{10}$$
$$= \frac{\sum_{m=1}^{M} p(y_k[n]|\mathcal{H}_m[n]=1)\Pr(\mathcal{H}_m[n]=1 \,|\, \mathcal{H}[n]=1)}{p(y_k[n] \,|\, \mathcal{H}[n]=0)} ,$$

where $\Pr(\mathcal{H}_m[n]=1|\mathcal{H}[n]=1)$ is estimated exploiting the approximation in Eq. (5). Such estimate is denoted as $\varphi_m[n]$:

$$\varphi_m[n] \triangleq \Pr(\mathcal{H}_m[n]=1|\mathcal{H}[n]=1)$$
$$= \frac{\Pr(\mathcal{H}_m[n]=1)}{\Pr(\mathcal{H}[n]=1)} \approx \frac{1-e^{-\widehat{\lambda}_m[n]\tau[n]}}{M-\sum_{m=1}^{M} e^{-\widehat{\lambda}_m[n]\tau[n]}} , \tag{11}$$

where $\widehat{\lambda}_m[n]$ denotes the latest available estimate of λ_m at the nth instant. Therefore, exploiting Eq. (9), we obtain the following ML Detector:

$$\Lambda_{k,n}(y_k[n]) = \sum_{m=1}^{M} \left(\varphi_m[n]\, a_{m,k}\, e^{b_{m,k}\, y_k^2[n]}\right) \underset{d_k[n]=0}{\overset{d_k[n]=1}{\gtrless}} 1 , \tag{12}$$

where

$$a_{m,k} = \sqrt{\frac{\sigma_{w,k}^2}{\sigma_{\xi,m}^2\, g^2(\boldsymbol{x}_k, \boldsymbol{h}_m) + \sigma_{w,k}^2}} , \tag{13}$$

$$b_{m,k} = \frac{1}{2}\left(\frac{1}{\sigma_{w,k}^2} - \frac{1}{\sigma_{\xi,m}^2\, g^2(\boldsymbol{x}_k, \boldsymbol{h}_m) + \sigma_{w,k}^2}\right) . \tag{14}$$

Since $\Lambda_{k,n}(y_k[n])$ in Eq. (12) is monotonically increasing in the variable $y_k^2[n]$, there exists a unique value $\gamma_k[n]$ such that:

$$\Lambda_{k,n}\left(\sqrt{\gamma_k[n]}\right) = 1 . \tag{15}$$

As a consequence, by the Karlin-Rubin Theorem, the test in Eq. (12) can be substituted with the following equivalent energy test [12]:

$$y_k^2[n] \underset{d_k[n]=0}{\overset{d_k[n]=1}{\gtrless}} \gamma_k[n] . \tag{16}$$

The task of numerically finding the value of $\gamma_k[n]$ that satisfies Eq. (15) is carried out by the PPC that continuously communicates the value of $\gamma_k[n]$ to the sensors.

For the kth sensor, the *probability of detection* ($P_{D,k,m}[n]$) when the mth item is in failed state and *probability of false alarm* ($P_{F,k}[n]$) with respect to the energy test in Eq. (16), at the nth instant, can be easily found as a consequence of Eq. (9) and are the following [13]:

$$P_{D,k,m}[n] \triangleq \Pr(d_k[n]=1|\mathcal{H}_m[n]=1) \tag{17}$$
$$= 2\mathcal{Q}\left(\sqrt{\gamma_k[n] \Big/ \left[\sigma_{\xi,m}^2\, g^2(\boldsymbol{x}_k, \boldsymbol{h}_m) + \sigma_{w,k}^2\right]}\right) ,$$

$$P_{F,k}[n] \triangleq \Pr(d_k[n]=1|\mathcal{H}[n]=0) \tag{18}$$
$$= 2\mathcal{Q}\left(\sqrt{\gamma_k[n] \,/\, \sigma_{w,k}^2}\right) .$$

B. Fusion Center Detection

The FC, at the nth instant, receives $\boldsymbol{d}[n]$ and, similarly to all sensors, performs a ML Detection as a consequence of the LRT based on Eqs. (8) and (9) [4]:

$$\Lambda_n^{\mathrm{FC}}(\boldsymbol{d}[n]) = \frac{\Pr(\boldsymbol{d}[n] \,|\, \mathcal{H}[n]=1)}{\Pr(\boldsymbol{d}[n] \,|\, \mathcal{H}[n]=0)}$$
$$= \sum_{m=1}^{M}\left\{\varphi_m[n]\prod_{k=1}^{K}\left[\left(\frac{P_{D,k,m}[n]}{P_{F,k}[n]}\right)^{d_k[n]} \times \left(\frac{1-P_{D,k,m}[n]}{1-P_{F,k}[n]}\right)^{1-d_k[n]}\right]\right\} \underset{\mathcal{D}[n]=0}{\overset{\mathcal{D}[n]=1}{\gtrless}} 1 , \tag{19}$$

where the values of $\varphi_m[n]$, $P_{D,k,m}[n]$'s, and $P_{F,k}[n]$'s are transmitted to the FC by the PPC.

Also for the FC it is possible to obtain the (FC) *probability of detection* ($Q_{D,m}[n]$) when the mth item is in failed state and the *probability of false alarm* ($Q_F[n]$) at the nth instant:

$$Q_{D,m}[n] \triangleq \Pr(\mathcal{D}[n]=1|\mathcal{H}_m[n]=1) \tag{20}$$
$$= \sum_{\boldsymbol{d}:\Lambda_n^{\mathrm{FC}}(\boldsymbol{d})\geq 1}\left\{\prod_{k=1}^{K}\left[P_{D,k,m}[n]^{d_k}\left(1-P_{D,k,m}[n]\right)^{1-d_k}\right]\right\} ,$$

$$Q_F[n] \triangleq \Pr(\mathcal{D}[n]=1|\mathcal{H}[n]=0) \tag{21}$$
$$= \sum_{\boldsymbol{d}:\,\Lambda_n^{\mathrm{FC}}(\boldsymbol{d})\geq 1}\left\{\prod_{k=1}^{K}\left[P_{F,k}[n]^{d_k}\left(1-P_{F,k}[n]\right)^{1-d_k}\right]\right\} .$$

C. Post-Processing Center Elaboration

1) Post-Processing Detection: The PPC has the main task of collecting $\mathcal{D}[n]$ and establishing whether an alarm should be raised. Unlike the local and global detection, the PPC *integrates the knowledge of the failure model*. Moreover, it takes advantage of all the values of $\mathcal{D}[j]$ where $j=1,\ldots,n$ (with $j=1$ is the first algorithm instance after the last inspection) to implement an effective quickest fault detection approach. For compactness, we define the (accumulated) vector $\boldsymbol{\mathcal{D}}[n] \triangleq \left[\mathcal{D}[1] \;\; \cdots \;\; \mathcal{D}[n]\right]^T$. In this case, the PPC performs a test equivalent to the following MAP Detector:

$$\Lambda_n^{\mathrm{PPC}}(\boldsymbol{\mathcal{D}}[n]) \triangleq \frac{\Pr(\boldsymbol{\mathcal{D}}[n]|\mathcal{H}[n]=1)}{\Pr(\boldsymbol{\mathcal{D}}[n]|\mathcal{H}[n]=0)} \underset{\widehat{\mathcal{H}}[n]=0}{\overset{\widehat{\mathcal{H}}[n]=1}{\gtrless}} \frac{\Pr(\mathcal{H}[n]=0)}{\Pr(\mathcal{H}[n]=1)} , \tag{22}$$

which is equivalent to the following test:

$$\mathcal{R}[n] \triangleq \Pr(\mathcal{H}[n]=1|\boldsymbol{\mathcal{D}}[n]) = \sum_{m=1}^{M} \Pr(\mathcal{H}_m[n]=1|\boldsymbol{\mathcal{D}}[n]) = \sum_{m=1}^{M} \mathcal{R}_m[n] \underset{\widehat{\mathcal{H}}[n]=0}{\overset{\widehat{\mathcal{H}}[n]=1}{\gtrless}} \frac{1}{2}\,, \tag{23}$$

where we exploited Eq. (5). By looking at the test in Eq. (23), it can be recognized that our approach is *optimal from a Bayesian viewpoint* (i.e. assuming the change point is a random variable, whose pdf originates from the reliability assumptions made in Sec. II-A) and corresponds to the Shiryaev decision statistic [14] with a threshold $(1/2)$ ensuring the minimization of a uniform Bayesian risk.

In the following, we detail how the expression of $\mathcal{R}_m[n]$ (for each m) can be updated in a recursive form based on the previous value $\mathcal{R}_m[n-1]$. First, capitalizing Bayes' Theorem and the conditional (i.e. given $\mathcal{H}_m[n]=1$) independence of FC decisions' $\mathcal{D}[1],\ldots,\mathcal{D}[n]$ in time, we obtain Eq. (27) for the mth item at the bottom of the page.

The latter expression depends on (i) the FC decision likelihood at time n ($\Pr(\mathcal{D}[n]|\mathcal{H}_m[n]=1)$) and (ii) the one-step prediction of the mth fault probability ($\Pr(\mathcal{H}_m[n]=1|\boldsymbol{\mathcal{D}}[n-1])$). The former likelihood is obtained as follows:

$$\Pr(\mathcal{D}[n]\,|\,\mathcal{H}_m[n]=1) = Q_{D,m}[n]^{\mathcal{D}[n]}\,(1-Q_{D,m}[n])^{(1-\mathcal{D}[n])} \tag{24}$$

$$\Pr(\mathcal{D}[n]\,|\,\mathcal{H}_m[n]=0) = Q_{F}[n]^{\mathcal{D}[n]}\,(1-Q_{F}[n])^{(1-\mathcal{D}[n])} \tag{25}$$

Conversely, the term $\Pr(\mathcal{H}_m[n]=1|\boldsymbol{\mathcal{D}}[n-1])$ can be rewritten as Eq. (28) at the bottom of the page. Such expression is obtained leveraging: (i) the independence of $\mathcal{H}_m[n]$ from $\boldsymbol{\mathcal{D}}[n-1]$ given $\mathcal{H}_m[n-1]$, and (ii) the inability of an item to self-repair (i.e. $\Pr(\mathcal{H}_m[n]=1|\mathcal{H}_m[n-1]=1)=1$). Furthermore, in Eq. (28) the term $\Pr(\mathcal{H}_m[n]=1|\mathcal{H}_m[n-1]=0)$ can be estimated explicitly using Eq. (3):

$$\Pr(\mathcal{H}_m[n]=1|\mathcal{H}_m[n-1]=0) = 1-e^{-\widehat{\lambda}_m[n]\Delta t}\,. \tag{26}$$

Combining the previous results, we obtain Eq. (29) at the bottom of the page as the final expression for $\mathcal{R}_m[n]$. Eq. (29) has been obtained in sequential form so that the PPC, at the nth instant, needs to store only the M values of $\mathcal{R}_m[n-1]$ and the value of $\mathcal{D}[n]$, instead of the n values present in $\boldsymbol{\mathcal{D}}[n]$. Eq. (29) requires initialization: still, it can be easily demonstrated that $\mathcal{R}_m[0]=0$.

2) Failure Rate Estimation: The exact failure rate of the generic mth item is unknown, however, literature can often provide an unbiased estimate ($\lambda_{m,0}$), as well as the corresponding variance (ν_m). However, literature data is typically obtained with a finite number of experiments on items that are not necessarily equal to those present in the system (or in the same operating conditions). Therefore, each λ_m is treated herein by the PPC as a random variable.

In detail, when an alarm is raised by the PPC, the system is shut down and an inspection is carried out to verify the condition of the system. If the mth item's ith failure is acknowledged, it is possible to compute an updated estimate of λ_m using $T_{m,i}$. As $T_{m,i}$ is not directly available, the assumption here is that $T_{m,j} \approx T^*_{m,j}$, which holds if $\varepsilon_{m,j} \ll \lambda_m$ for all $j=1,\ldots,i$ (i.e. the expected item lifetime is much higher than the delay accumulated by the system to detect the fault). By defining the vector $\boldsymbol{T}_m[i] \triangleq \begin{bmatrix} T_{m,1} & \cdots & T_{m,i} \end{bmatrix}^T$, the PPC computes the following *Minimum Mean Square Error Estimator* (MMSE) for the mth item:

$$\widehat{\lambda}_{m,i} = \mathbb{E}\{\lambda_m|\boldsymbol{T}_m[i]\}\,. \tag{30}$$

In order to evaluate such expectation, the PPC needs to compute the (posterior) pdf of $\lambda_m\,|\,\boldsymbol{T}_m[i]$. Since $T_{m,j} \sim \text{Exp}(\lambda_m)$ for all $j=1,\ldots,i$, we embody the prior knowledge about the mth item lifetime via $\lambda_m \sim \text{Gamma}(\alpha_{m,0},\,\beta_{m,0})$, where $\alpha_{m,0} \triangleq (\lambda^2_{m,0}\,/\,\nu_m)$ and $\beta_{m,0} \triangleq (\lambda_{m.0}\,/\,\nu_m)$ are obtained based on the available literature values. Our choice leverages the Gamma pdf since it is known to be the *conjugate prior* of the Exponential pdf [15]. Capitalizing the advantages of using a conjugate prior, it is not difficult to show that $\lambda_m\,|\,\boldsymbol{T}_m[i] \sim \text{Gamma}(\alpha_{m,i},\,\beta_{m,i})$, where the Gamma parameters can be computed *recursively* by the PPC as $\alpha_{m,i} = (\alpha_{m,i-1}+1)$ and $\beta_{m,i} = (\beta_{m,i-1}+T_{m,i})$, respectively.

Once the (Gamma) posterior pdf of $\lambda_m\,|\,\boldsymbol{T}_m[i]$ is calculated, the corresponding MMSE estimator after ith failure is obtained by exploiting standard properties of Gamma distribution:

$$\widehat{\lambda}_{m,i} = \alpha_{m,i}\,/\,\beta_{m,i} \tag{31}$$

Clearly, at generic time n, the latest available estimate of λ_m is obtained as $\widehat{\lambda}_{m,\mathcal{S}_m} \rightarrow \widehat{\lambda}_m[n]$, assuming $\mathcal{S}_m$ failures for mth item have been reported up to time $(n-1)$.

$$\mathcal{R}_m[n] = \frac{\Pr(\mathcal{D}[n]|\mathcal{H}_m[n]=1)\,\Pr(\mathcal{H}_m[n]=1|\boldsymbol{\mathcal{D}}[n-1])}{\Pr(\mathcal{D}[n]|\mathcal{H}_m[n]=1)\,\Pr(\mathcal{H}_m[n]=1|\boldsymbol{\mathcal{D}}[n-1]) + \Pr(\mathcal{D}[n]|\mathcal{H}_m[n]=0)\,[1-\Pr(\mathcal{H}_m[n]=1|\boldsymbol{\mathcal{D}}[n-1])]} \tag{27}$$

$$\Pr(\mathcal{H}_m[n]=1|\boldsymbol{\mathcal{D}}[n-1]) = \mathcal{R}_m[n-1] + \Pr(\mathcal{H}_m[n]=1|\mathcal{H}_m[n-1]=0)(1-\mathcal{R}_m[n-1]) \tag{28}$$

$$\mathcal{R}_m[n] = \frac{\Pr(\mathcal{D}[n]|\mathcal{H}_m[n]=1)\left[1-e^{-\widehat{\lambda}_m[n]\Delta t}(1-\mathcal{R}_m[n-1])\right]}{\Pr(\mathcal{D}[n]|\mathcal{H}_m[n]=1)\left[1-e^{-\widehat{\lambda}_m[n]\Delta t}(1-\mathcal{R}_m[n-1])\right] + \Pr(\mathcal{D}[n]|\mathcal{H}_m[n]=0)\left[e^{-\widehat{\lambda}_m[n]\Delta t}(1-\mathcal{R}_m[n-1])\right]} \tag{29}$$

3) Parameters Calculation and Communication: At the beginning of a next $(n+1)$ step, the PPC updates (if needed) the estimates of the failure rates of the appropriate items to obtain $\widehat{\lambda}_m[n+1]$, and then calculates the values of $\varphi_m[n+1]$'s via Eq. (11). Next, it computes and transmits the values of $\gamma_k[n+1]$'s to the respective sensors to be used for the next energy test (see Eq. (16)). This consists of finding the $\gamma_k[n+1]$ that solves Eq. (15) for the $(n+1)$th instant for all sensors:

$$\sum_{m=1}^{M} \left(\varphi_m[n+1]\, a_{m,k}\, e^{b_{m,k}\, \gamma_k[n+1]} \right) = 1 \,. \tag{32}$$

The left-hand side of Eq. (32) is a smooth, convex, and increasing function in the variable $\gamma_k[n+1]$, therefore convergence is guaranteed from any starting point $\gamma_k[n+1]^{(0)}$ when solving it using the *Newton-Raphson method* [16]:

$$\gamma_k[n+1]^{(q+1)} \tag{33}$$
$$= \gamma_k[n+1]^{(q)} - \frac{\sum_{m=1}^{M} \left(\varphi_m[n+1]\, a_{m,k}\, e^{b_{m,k}\, \gamma_k[n+1]^{(q)}} \right) - 1}{\sum_{m=1}^{M} \left(\varphi_m[n+1]\, a_{m,k}\, b_{m,k}\, e^{b_{m,k}\, \gamma_k[n+1]^{(q)}} \right)},$$

where q denotes the iteration index.

Once obtained the thresholds, the PPC computes the values of $P_{D,k,m}[n+1]$'s and $P_{F,k}[n+1]$'s via Eqs. (17) and (18) and transmits them to the FC together with the values of $\varphi_m[n+1]$'s. This allows the FC to evaluate $\Lambda_{n+1}^{\mathrm{FC}}(\boldsymbol{d}[n+1])$ via Eq. (19). As a last step, the PPC computes the values of $Q_{D,m}[n+1]$'s and $Q_F[n+1]$ via Eqs. (20) and (21) to be used by the PPC itself in the (recursive) computation of $\Lambda_{n+1}^{\mathrm{PPC}}(\boldsymbol{\mathcal{D}}[n+1])$ via Eqs. (24) and (25). Note that Eqs. (20) and (21) can be calculated exactly with a finite number of operations since the number of possible outcomes of $\Lambda_n^{\mathrm{FC}}(\boldsymbol{d})$ is 2^K.

IV. Case Study – Goliat FPSO

The Goliat FPSO is an oil production platform located in the Barents Sea. Such platform relies on a Subsea Production System made of eight templates installed on the seabed[2]. As a consequence, oil leaks happen in deep waters making the detection even more challenging. Moreover, because of the high depths, the inspections must be performed by *remotely operated vehicles* with high costs making the reduction of false alarms of great importance [18]. On the other hand, strict environmental regulations are enforced on companies operating offshore, asking for quick detection of the spills [19]. A characteristic of oil leaks is their related acoustic signal that can be sensed via passive acoustic sensors [20], [21]. Therefore, each template has its manifold monitored by $K = 3$ *passive acoustic sensors* measuring the sound pressure as part of the leak detection system [22], [23]. An analysis recognized $M = 20$ *critical items* assumed to be at the same height as the sensors, as shown in Fig. 3. The previously described algorithm is assumed to be implemented over the system in place to verify the performance. The AAF used for this case study is the following:

$$g(\boldsymbol{x}_k, \boldsymbol{\theta}) = \sqrt{\left(\frac{\ell_{\mathrm{ref}}}{\|\boldsymbol{x}_k - \boldsymbol{\theta}\|} \right)^{k_{\mathrm{sc}}} 10^{(\ell_{\mathrm{ref}} - \|\boldsymbol{x}_k - \boldsymbol{\theta}\|)\alpha 10^{-4}}} \,, \tag{34}$$

where ℓ_{ref} and $\|\boldsymbol{s}_k - \boldsymbol{\theta}\|$ are in meters, the seawater absorption coefficient α is in [dB/km], and k_{sc} is the spreading coefficient. The value of α has been computed using *Francois & Garrison equation* [24], [25], with the underwater speed of sound obtained via the *Chen & Millero equation* [26] with parameters in Tab. I.

The proposed algorithm is compared with a WSN without the PPC, meaning that the sensors will perform the energy test in Eq. (16), and the FC, after collecting $\boldsymbol{d}[n]$, will provide the final decision on whether an oil spill is occurring at time n via Eq. (19), i.e. without exploiting the temporal dimension. It can be shown that the latter approach corresponds to a quickest oil spill detection based on the classical *Shewhart chart* [14]. However, the absence of a PPC removes the possibility to update parameters and transmit them to the sensors and the FC. Therefore, the values of φ_m's are taken as constants:

$$\varphi_m = \Pr(\mathcal{H}_m = 1 | \mathcal{H} = 1) \approx \lambda_{m,0} \Big/ \sum_{m=1}^{M} \lambda_{m,0} \,, \tag{35}$$

TABLE I
Parameters used for the simulation

Parameter	Value	Note / Reference
Ref. Frequency	2.5 kHz	[27]
Temperature	3.8 °C	[28]
Salinity	35 ‰	[28]
Depth	350 m	[22]
pH	8	[29]
k_{sc}	1.5	[30]
ℓ_{ref}	1 m	–
Simulated time	15 yr	[31]
Δt	15 min	–
$SNR_{m,k}$	10/15/20/25 dB	$\forall\, m, k$

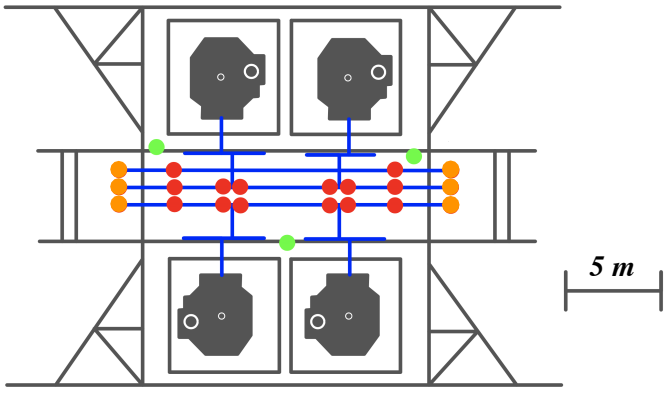

Fig. 3. Goliat's subsea template: the gray elements are the structure, the blue lines are the manifold, the green dots are the passive acoustic sensors, the red dots are the valves, and the orange dots are the connectors.

[2]For further details on subsea production systems, please see [17].

TABLE II
LITERATURE VALUES OF FAILURE RATES OF SUBSEA ITEMS

Item Category	$\lambda_{m,0}$ [yr^{-1}]	ν_m [yr^{-2}]
Valve, process isolation (manifold)	7.3000×10^{-3}	7.0715×10^{-5}
Connector (manifold)	9.5812×10^{-4}	2.4649×10^{-6}

where the literature values of the failure rates $\lambda_{m,0}$'s are used as parameter estimates. Eq. (35) was obtained exploiting the properties of a merged Poisson process. The time-independent property of the values of φ_m's also reflects on the values of γ_k's, $P_{D,k,m}$'s, and $P_{F,k}$'s which no longer need to be updated.

Numerical results have been obtained via simulations with 200 Monte Carlo runs using the software MATLAB. In detail, each run simulates the life of the Goliat FPSO (assuming negligible inspection time and not accounting for maintenance time). The simulated time, the value of Δt, and the different values of $SNR_{m,k} \triangleq \sigma^2_{\xi,m}/\sigma^2_{w,k}$ are in Tab. I. At each run, for each item, a new realization of the M Poisson processes and their corresponding failure rates is generated, where the λ_m's are originated from a Gamma distribution with moments from Tab. II which reports the literature values retrieved from *OREDA Handbook* [32].

The main results are in Tab. III where the aggregated average of the values of $\varepsilon_{m,j}$ (and the corresponding number of collected samples), the *Fault Rate* $P_1 = \Pr(\mathcal{H}[n] = 1)$, and the *False Positive Rate* $P_{10} = \Pr\left(\widehat{\mathcal{H}}[n] = 1 | \mathcal{H}[n] = 0\right)$ are displayed. Also, the *True Positive Rate* is defined as $P_{11}^{(N)} = \Pr\left(\bigcup_{j=0}^{N-1} \left\{\widehat{\mathcal{H}}[n+j] = 1\right\} | \, \mathcal{H}[n] = 1\right)$, where $(N-1)\Delta t$ is the allowed detection delay, with $N \geq 1$ as the number of collected samples. Fig. 4 shows the True Positive Rate as a function of N for different relevant SNR spill values. Moreover, the *Error Rate* (P_e) as a function of N is reported in Fig. 5 and is calculated as follows:

$$P_e^{(N)} = P_{10}(1 - P_1) + \left(1 - P_{11}^{(N)}\right) P_1 \, . \tag{36}$$

From Tab. III we notice the significant difference in behavior between the two architectures. When the Reliability-Based Fusion is implemented, the WSN performs a temporal integration with a consequent higher number of measurements needed by the PPC to declare the presence of a leak as it can be seen from the values of $\varepsilon_{m,j}$ which, on average, are higher when the PPC is implemented. This originates two opposing effects: on one hand, without the PPC, the WSN reaches a higher value of True Positive Rate for a given value of N as shown in Fig. 4, on the other hand, the temporal integration of the PPC allows the WSN to dramatically decrease the value of False Positive Rate of *five orders of magnitude*. It is also important to notice that Fault Rate will be higher in case the PPC is implemented as a consequence of the tendency of showing higher values of $\varepsilon_{m,j}$. Thus, it is evident the need to use the Error Rate as a metric in order to better evaluate

TABLE III
SIMULATION RESULTS

	$SNR_{m,k}$ [dB]	With PPC	Without PPC
Average no. failures	–	1.5350	1.5350
Average $\varepsilon_{m,j}$ (no. samples)	10	32 hr 23 min (130.56)	5 min 46 s (1.38)
	15	11 hr 40 min (47.69)	4 min 50 s (1.32)
	20	5 hr 13 min (21.85)	2 min 59 s (1.20)
	25	2 hr 30 min (11)	4 min 3 s (1.27)
Fault Rate	10	3.8104×10^{-4}	4.0402×10^{-6}
	15	1.3918×10^{-4}	3.8596×10^{-6}
	20	6.3779×10^{-5}	3.4984×10^{-6}
	25	3.2103×10^{-5}	3.7075×10^{-6}
False Positive Rate	10	5.0405×10^{-6}	5.4641×10^{-1}
	15	4.9821×10^{-6}	4.4817×10^{-1}
	20	3.3560×10^{-6}	3.3096×10^{-1}
	25	2.1866×10^{-6}	1.4331×10^{-1}

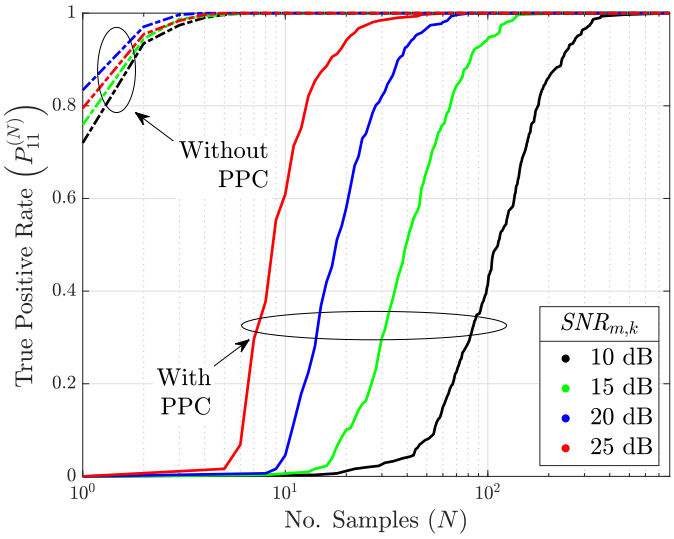

Fig. 4. True Positive Rate as a function of the number of collected samples N.

the trade-off between quick detection (fast spill detection) and a high number of false alarms (high inspection costs).

Finally, Fig. 5 shows the Error Rate versus N. In absence of PPC, the Error Rate is found to be greater than 0.1 without significant changes with N, whereas such value decreases as $SNR_{m,k}$ increases. Indeed, a fault detection mechanism not capitalizing time integration incurs in a high False Positive Rate (see Tab. III), which represents the dominating term in Eq. (36). On the contrary, when the PPC is implemented, the Error Rate starts from a value $\in [10^{-5}, 10^{-3}]$ and, as N increases, decreases by settling around a value $\in [10^{-6}, 10^{-5}]$. Such decrease in the Error Rate with N (when the PPC is present) is due to the corresponding increase of True Positive Rate made possible by the temporal integration performed by the PPC. Clearly, a higher value of $SNR_{m,k}$ lowers the Error Rate regardless of the employed architecture. Hence, using the

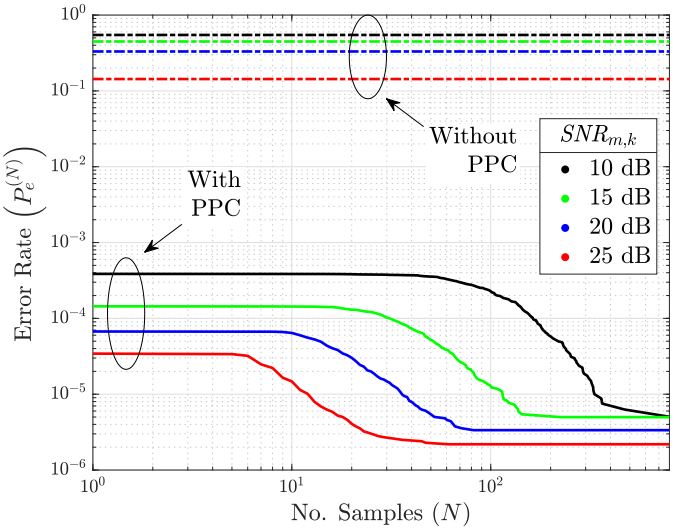

Fig. 5. Error Rate as a function of the number of collected samples N.

Error Rate as a metric, our Reliability-Based Fusion algorithm evidently outperforms the same architecture lacking the PPC.

V. Conclusions and Future Directions

In this work, we tackled quickest detection of faults within an Oil and Gas subsea production system, by means of spatio-temporal decision fusion approach. The sensor network collectively monitors the state of different pieces of equipment and reports their decisions to a FC based on individual LRTs. Herein, a spatial aggregation is performed, based on a global (per-sample) LRT and a global decision is performed. Such decisions are then aggregated in time by a PPC, which performs quickest detection of the system state according to a Bayesian criterion and exploits statistical distributions about the change time driven by datasheet reliability metrics. Results have highlighted the benefit in terms of Error Rate of a reliability-based algorithm with respect to an architecture that does not include the knowledge of the reliability features of the monitored system in its design. *Future directions* of research will include: (a) reliability-aided quickest fault detection in the presence of unknown parameters, (b) lossy reporting channels and (c) considering more complex reliability models.

References

[1] T. Sahoo, *Process Plants - Shutdown and Turnaround Management*. Boca Raton (FL), USA: Taylor & Francis Group, 2014.

[2] Y. Song, "Underwater Acoustic Sensor Networks With Cost Efficiency for Internet of Underwater Things," *IEEE Trans. Ind. Electron.*, vol. 68, no. 2, pp. 1707–1716, Feb. 2021.

[3] N. Paltrinieri, G. Landucci, and P. Salvo Rossi, "Real-Time Data for Risk Assessment in the Offshore Oil and Gas Industry," in *Proc. Int. Conf. Omae*, Trondheim, Norway, 2017.

[4] D. Ciuonzo and P. Salvo Rossi, "Distributed detection of a non-cooperative target via generalized locally-optimum approaches," *Inf. Fusion*, vol. 36, pp. 261–274, Jul. 2017.

[5] N. Sriranga, K. G. Nagananda, R. S. Blum, A. Saucan, and P. K. Varshney, "Energy-Efficient Decision Fusion for Distributed Detection in Wireless Sensor Networks," in *Proc. Int. Conf. Inf. Fusion*, Cambridge, UK, 2018, pp. 1541–1547.

[6] S. Sen, N. S. V. Rao, C. Q. Wu, M. L. Berry, K. M. Grieme, R. R. Brooks, and G. Cordone, "Performance analysis of Wald-statistic based network detection methods for radiation sources," in *Proc. Int. Conf. Inf. Fusion*, Heidelberg, Germany, 2016, pp. 820–827.

[7] G. Tabella, N. Paltrinieri, V. Cozzani, and P. Salvo Rossi, "Subsea oil spill risk management based on sensor networks," *Chem. Eng. Trans.*, vol. 82, pp. 199–204, Oct. 2020.

[8] M. Bucelli, I. B. Utne, P. Salvo Rossi, and N. Paltrinieri, "A system engineering approach to subsea spill risk management," *Saf. Sci.*, vol. 123, Mar. 2020.

[9] G. Tabella, N. Paltrinieri, V. Cozzani, and P. Salvo Rossi, "Data fusion for subsea oil spill detection through wireless sensor networks," in *Proc. IEEE Sensors*, Rotterdam, The Netherlands, 2020.

[10] ——, "Wireless sensor networks for detection and localization of subsea oil leakages," *IEEE Sens. J.*, vol. 21, no. 9, pp. 10 890–10 904, May 2021.

[11] J. Pan and J. McElhannon, "Future edge cloud and edge computing for internet of things applications," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 439–449, 2017.

[12] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Pacific Groove (CA), USA: Thomson Learning, 2002.

[13] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, 1st ed., ser. (Prentice Hall Signal Processing Series). Upper Saddle River (NJ), USA: Prentice-Hall, 1998.

[14] L. Xie, S. Zou, Y. Xie, and V. V. Veeravalli, "Sequential (quickest) change detection: Classical results and new directions," *IEEE Journal on Selected Areas in Information Theory*, 2021.

[15] M. Rausand and S. Haugen, *Risk Assessment: Theory, Methods, and Applications*, 1st ed. Hoboken (NJ), USA: John Wiley & Sons, 2011.

[16] D. Kincaid and W. Cheney, *Numerical Analysis: Mathematics of Scientific Computing*, 3rd ed. Providence (RI), USA: AMS, 2002.

[17] Y. Bai and Q. Bai, *Subsea Engineering Handbook*. Houston (TX), USA: Elsevier, 2012.

[18] C. Mai, S. Pedersen, L. Hansen, K. L. Jepsen, and Zhenyu Yang, "Subsea infrastructure inspection: A review study," in *IEEE Int. Conf. Underwater Syst. Technol., Theory Appl. (USYS)*, Penang, Malaysia, Dec. 2016, pp. 71–76.

[19] DNV-GL, "Recommended practice RP-F302 offshore leak detection," Oslo, Norway, Apr. 2016.

[20] H. V. Fuchs and R. Riehle, "Ten years of experience with leak detection by acoustic signal analysis," *Appl. Acoust.*, vol. 33, no. 1, pp. 1–19, 1991.

[21] M. A. Adegboye, W. K. Fung, and A. Karnik, "Recent advances in pipeline monitoring and oil leakage detection technologies: Principles and approaches," *Sensors*, vol. 19, no. 11, p. 2548, Jun. 2019.

[22] E. Bjørnbom, "Goliat – Leak detection and monitoring from template to satellite," 2011.

[23] E. Røsby, "Goliat development project - Subsea leak detection design," Aker Solutions, 2011.

[24] R. E. Francois and G. R. Garrison, "Sound absorption based on ocean measurements: Part I: Pure water and magnesium sulfate contributions," *J. Acoust. Soc. Am.*, vol. 72, no. 3, pp. 896–907, 1982.

[25] ——, "Sound absorption based on ocean measurements: Part II: Boric acid contribution and equation for total absorption," *J. Acoust. Soc. Am.*, vol. 72, no. 6, pp. 1879–1890, Dec. 1982.

[26] G. S. K. Wong and S. Zhu, "Speed of sound in seawater as a function of salinity, temperature, and pressure," *J. Acoust. Soc. Am.*, vol. 97, no. 3, pp. 1732–1736, Mar. 1995.

[27] E. G. Eckert, J. W. Maresca, R. W. Hillger, and J. J. Yezzi, "Location of leaks in pressurized petroleum pipelines by means of passive-acoustic sensing methods," in *Leak Detection for Underground Storage Tanks*. West Conshohocken (PA), USA: ASTM Int., 1993, pp. 53–69.

[28] Institute of Marine Research, "Mareano," 2021.

[29] A. A. Vetrov and E. A. Romankevich, *Carbon Cycle in the Russian Arctic Seas*. Berlin, Germany: Springer, 2004.

[30] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," in *Proc. Int. Workshop Underwater Networks (WUWNet)*, Oct. 2006, pp. 34–43.

[31] Vår Energi, "Goliat Barrier Status Panel," 2016.

[32] SINTEF, *OREDA Offshore Reliability Data Handbook*, 4th ed. OREDA Participants, 2002.