Sensor Fusion for Detection and Localization of Carbon Dioxide Releases for Industry 4.0

Gianluca Tabella*†, Yuri Di Martino, Domenico Ciuonzo‡, Nicola Paltrinieri§, Xiaodong Wang†, and Pierluigi Salvo Rossi*¶

*Dept. Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway

†Dept. Electrical Engineering, Columbia University, New York, NY, USA

‡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; yuri.di.martino@gmail.com; domenico.ciuonzo@unina.it; nicola.paltrinieri@ntnu.no; wangx@ee.columbia.edu; salvorossi@ieee.org

***Abstract*—This work tackles the distributed detection & localization of carbon dioxide (CO_2) release from storage tanks caused by the opening of pressure relief devices via inexpensive sensor devices in an industrial context. A realistic model of the dispersion is put forward in this paper. Both full-precision and rate-limited setups for sensors are considered, and fusion rules capitalizing the dispersion model are derived. Simulations analyze the performance trends with realistic system parameters (e.g. wind direction).**

***Index Terms*—Carbon Dioxide (CO_2), Decision Fusion, Detection, Industry 4.0, Internet of Things, Localization, Wireless Sensor Networks.**

I. Introduction

The last decades have seen the growth of Wireless Sensor Networks (WSNs) due to their collective, cost-effective, and successful use in industrial & environmental monitoring applications [1]. This surge has become even more pronounced with the rise of the Internet of Things (IoT) paradigm. In particular, the discovery of *harmful events* has received large attention: relevant scenarios include (i) counter-terrorism, (ii) safety & security in Industry 4.0, and (iii) environmental protection [2].

In the above context, the associated inference problems are "early" detection of an unknown source and its precise localization [3], [4]. In this context, most of the existing works only assume a Gaussian plume point source model based on diffusion/advection processes, e.g. with application to dispersion of biochemical moving sources [5], [6], localization of atmospheric pollutants [7] and release of light gases [8]. On the contrary, *carbon dioxide* (CO_2) is a (heavy) gas whose density, at atmospheric temperature and pressure, is about 1.5 larger than the air density and is present in atmosphere at an average concentration around 400 ppm, as of today. Nowadays, CO_2 finds several applications at domestic and industrial levels [9], [10]. Unluckily, when CO_2 is stored, it is possible that accidental releases occur with the main danger of asphyxiation. Being a heavy gas, CO_2 *does not adhere* to neutral or positively-buoyant dispersion behavior.

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).

For bulk storage, CO_2 is typically stored as liquid in insulated tanks[1] (see Fig. 1), usually equipped with systems to limit the internal pressure, namely *pressure relief devices* (PRDs). These can be safety valves, rupture disks, or their combinations. PRDs are designed in accordance to international or national standards to protect the vessel when the internal pressure exceeds the *maximum allowable working pressure* (MAWP). The causes of overpressure may be several, ranging from process upsets to external fires. In any of these cases, the PRD must release the flow rate necessary to avoid dangerous pressure build-up inside the tank. In such cases, however, the consequences of PRD activation can still be harmful to human life and accurate detection of these critical events should be performed leveraging WSNs.

To this end, an *industrial IoT* setup with inexpensive sensors and the possibility of leveraging collective (cloud-based) analytics to obtain improved performance (and global awareness of the monitored plant), represents an enabler for this problem. However, due to their stringent bandwidth and energy constraints needed to ensure long-lasting lifetime of IoT nodes, sensors are usually constrained to send extremely-compressed versions of their measurements to a Fusion Center (FC). For such a reason, the localization of diffusive sources via WSNs has shifted toward the adoption of binary sensors [12], [13].

Accordingly, the *contributions* of this work are as follows. We model the release of CO_2 from PRDs via a set of analytical relationships desumed from the well-known *Britter & McQuaid* (B&M) empirical model, which overcomes the usual (manual) nomogram inspection. Also, our formulation accounts for the unavoidable fluctuations in the concentration. The sensors measure the concentration at their location and report only one bit to the FC, targeting an industrial IoT setup with cheap small-battery sensors. Since the activated PRD is unknown, the FC is in charge of performing decision fusion by tackling a composite hypothesis testing. For the mentioned reason, a generalized likelihood ratio (GLR)-based fusion rule

[1]Storage temperature is below ambient temperature, typically $\in [-30, -20]$°C with corresponding pressures of $\in [14.3, 19.7]$ bar [10], [11].

is devised [14] and compared with a GLR counterpart based on full-precision measurements and the Counting Rule (CR) for the *detection task*. Once a PRD opening is detected, we also address the *localization task* to infer its position accurately, with the aim of speeding up maintenance operations (and thus diminish associated costs). In such a context, the raw/one-bit Maximum Likelihood Estimator (MLE), the centroid and Center of the Minimum Enclosing Circle (CMEC) estimators are compared. Detection & localization approaches are also compared in terms of the complexity involved. Simulation results highlight the need for including a realistic CO_2 release model within the design of fusion rules in *both tasks*.

The rest of the paper is organized as follows. Sec. II describes the system model considered, whereas Sec. III and IV introduce the proposed decision fusion approach for detection, while Sec. V focuses on localization strategies. Sec. VI gives an overview of the computational complexity of the proposed algorithms. Our approach is then numerically validated on a real case study in Sec. VII. Finally, Sec. VIII ends the paper with some research prospects.

Notation – Bold letters denote vectors; $(\cdot)^T$, and $\|\cdot\|$ denote transpose and Euclidean norm operators, respectively; $\hat{a}$, $\mathbb{E}(a)$, $\mathbb{V}\text{ar}(a)$, $\mathbb{E}(a|b)$, and $\mathbb{V}\text{ar}(a|b)$ denote an estimate of the random variable a, its expectation, its variance, its conditional expectation given the random variable b and conditional variance given 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; $F_a(\cdot)$ is the cumulative distribution function (cdf) of the random variable a and $F_{a|b}(\cdot)$ is its conditional counterpart given the random variable b; Gamma $(\alpha,\ \beta)$ denotes a Gamma distribution with shape α and rate β; $\mathcal{B}(p)$ denotes a Bernoulli distribution with parameter p; $\Gamma(\cdot)$ is the Gamma function; the symbol $\sim$ (resp. $\overset{\text{approx.}}{\sim}$) means "distributed as" (resp. "approximately distributed as"); finally $\mathcal{O}(\cdot)$ denotes the big O notation.

II. System Model

A. Wireless Sensor Network Model

The examined industrial facility consists of M vessels containing a heavy gas and their respective PRDs, where the mth PRD is located at (two-dimensional) position $\boldsymbol{h}_m = \begin{bmatrix} h_{m1} & h_{m2} \end{bmatrix}^T$. The plant is monitored by K concentration sensors with the kth sensor in position $\boldsymbol{x}_k = \begin{bmatrix} x_{k1} & x_{k2} \end{bmatrix}^T$. The K sensors individually assess the absence ($\mathcal{H}_0$) or presence ($\mathcal{H}_1$) of a gas dispersion by measuring the local gas concentration y_k and reporting their local decision $d_k = i$, if $\mathcal{H}_i$ is declared as reported in Fig. 2. Binary decisions are spectrally-efficient, as only 1-bit communication is required on the communication channel between the sensor and the FC, as well as being energy-efficient when OOK is employed [2], [15]. The vector of local decisions $\boldsymbol{d} = \begin{bmatrix} d_1 & \cdots & d_K \end{bmatrix}^T$ is acquired by the FC that processes it and takes a global decision $\widehat{\mathcal{H}} \in \{\mathcal{H}_0, \mathcal{H}_1\}$. When $\widehat{\mathcal{H}} = \mathcal{H}_1$, the FC also provides an estimate $\widehat{m} \in \{1, \ldots, M\}$ of the PRD declared as the source.

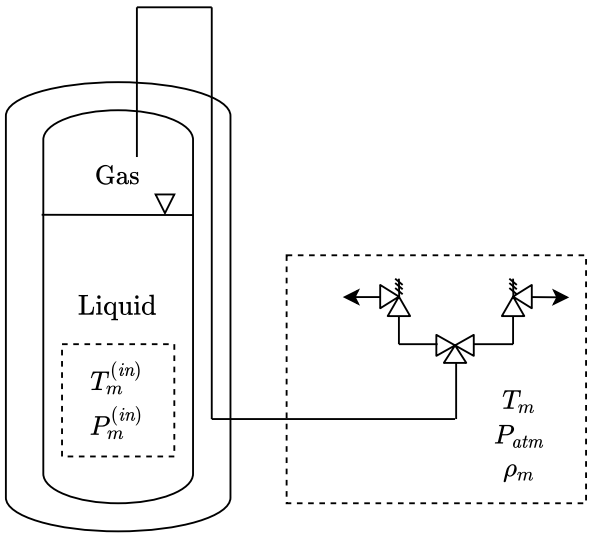


Fig. 1: Scheme of the tank and its PRD with the corresponding thermodynamic conditions.

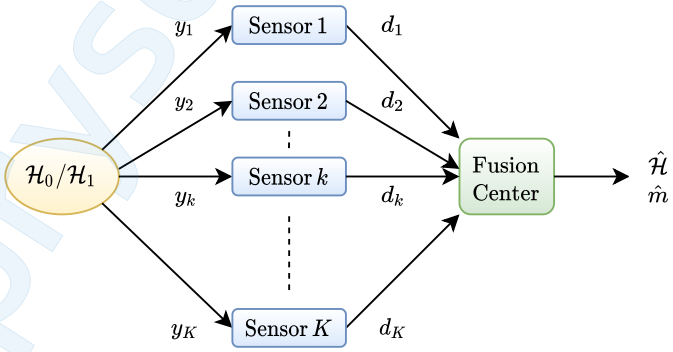


Fig. 2: Wireless Sensor Network Architecture.

As a comparative tool, the WSN is also examined in the case in which the FC acquires full-precision measurements $\boldsymbol{y} = \begin{bmatrix} y_1 & \cdots & y_K \end{bmatrix}^T$ from the sensors.

B. Dispersion Model

The heavy gas dispersion model here used is based on the well-known *Britter & McQuaid Model* (B&M) for continuous releases [16]–[21]. Such method, however, is based on the reading of a nomogram which prevents its utilization on a real-time basis by the FC. To this end, in this work we present a set of equations substituting the nomogram whose output is the value of the *average molar fraction concentration* at the kth sensor when the mth PRD is open, namely $c_{k,m}$ ($0 \leq c_{k,m} \leq 1$). More specifically, we provide the following map $\mathcal{F}$:

$$c_{k,m} = \mathcal{F}\left(\boldsymbol{x}_k, \boldsymbol{h}_m, T_m, \rho_m, c_m, \dot{V}_m, D_m, T_{atm}, \rho_{air}, u, \varphi\right), \tag{1}$$

where T_m, ρ_m, c_m, and $\dot{V}_m$ are the temperature, density, concentration, and volumetric flow rate (respectively) of CO_2 at release condition from the mth PRD, whose diameter is

denoted with D_m; T_{atm} is the atmospheric temperature; ρ_{air} is the density of air at T_{atm}; finally u and φ are the wind speed at a height of 10 meters[2] and its direction[3] (respectively). We now detail the mapping in Eq. (1) via the constituting set of relationships (and corresponding assumptions) reported in what follows.

1) Thermodynamic Properties: Inside the tank corresponding to the mth valve, there exists a CO_2 liquid-vapor equilibrium at a certain pressure $P_m^{(in)} > P_{atm}$ and the corresponding saturation temperature $T_m^{(in)} = T_{sat}\left(P_m^{(in)}\right)$. When $P_m^{(in)}$ reaches the set pressure, the device opens, releasing the gas phase in atmosphere. At this point a *Joule-Thompson process* occurs with a consequent (isenthalpic) expansion and cooling of the gas[4]. At release condition, the gas will be at temperature T_m, pressure P_{atm}, and density ρ_m. The values of T_m and ρ_m can be obtained through an appropriate equation of state (EOS) using $T_m^{(in)}$ and $P_m^{(in)}$ as inputs (see Fig. 1).

2) Applicability Criteria: The B&M model is meant for continuous release of heavy gases, so it is vital that the following criterion is met to consider the release "dense enough":

$$\frac{{g'_m}^2 \dot{V}_m^{0.5}}{u^{2.5}} \geq 3.375 \times 10^{-3} \,, \tag{2}$$

where $g'_m = g\left(\rho_m - \rho_{air}\right)/\rho_{air}$ and g is the gravitational acceleration.

Some limitations potentially affecting the accuracy of the model are:

- The release is assumed to be at ground level;
- The calculated concentrations are at ground level;
- The concentration in the cross-sectional area of the gas plume is assumed uniform;
- Jet due to a high-velocity release is not modeled;
- Obstacles are not modeled;
- Concentrations at low distances from the source have higher prediction error.

3) Change of Coordinates: All sensor positions must go through the change of coordinates represented in Fig. 3. This is because B&M centers the coordinate system at the source point with the first coordinate pointing downwind. The following rototranslation assumes that the vectors $\boldsymbol{x}_k$ and $\boldsymbol{h}_m$ are obtained from a map following the north-up standard map orientation. The following equation describes how to obtain this change when the mth PRD is open:

$$\boldsymbol{x}_k^{(m)} = \begin{bmatrix} -\sin\varphi & -\cos\varphi \\ \cos\varphi & -\sin\varphi \end{bmatrix} \left(\boldsymbol{x}_k - \boldsymbol{h}_m\right) \,, \tag{3}$$

4) Dimensionless Quantities: Two quantities relative to the mth PRD must be calculated to perform B&M:

$$L_m = \frac{\dot{V}_m g'_m}{u^3} \,, \quad \phi_m = \left(\frac{{g'_m}^2 \dot{V}_m}{u^5}\right)^{0.2} \,. \tag{4}$$

[2]If wind speed is available at a different height, several conversion methods are available [19].

[3]Wind blowing from north: 0° (360°), east: 90°, south: 180°, west: 270°.

[4]We neglect possible formation of liquid or solid during this transformation.

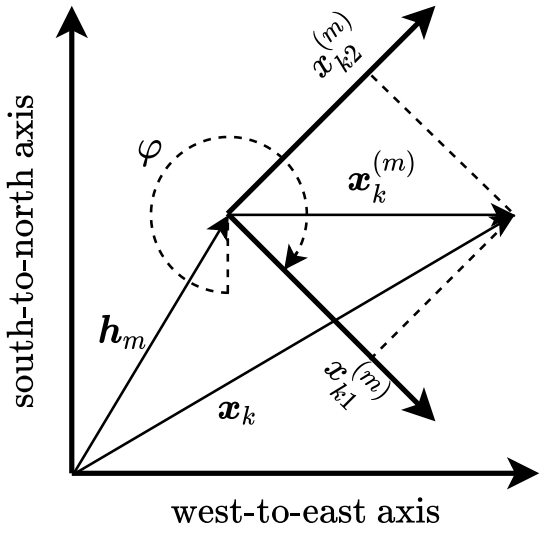


Fig. 3: Representation of the change of coordinates.

5) Concentration Ratio Calculation: The following procedure is intended to provide the concentration ratio in the whole monitored area:

$$\frac{c_{k,m}}{c_m} = \begin{cases} f_{k,m}, & x_{k1}^{(m)} > 0 \wedge \left|x_{k2}^{(m)}\right| \leq R_{W_{k,m}} \\ 1, & -x_{U_m} \leq x_{k1}^{(m)} \leq 0 \wedge \left|x_{k2}^{(m)}\right| \leq R_{U_{k,m}} \\ 0, & \text{otherwise} \end{cases} \,, \tag{5}$$

where x_{U_m} is the upwind distance, $R_{W_{k,m}}$ is the downwind radius, and $R_{U_{k,m}}$ is the upwind radius:

$$x_{U_m} = \frac{D_m}{2} + 2L_m \,, \tag{6}$$

$$R_{W_{k,m}} = R_{0_m} + 2.5 L_m^{1/3}\left(x_{k1}^{(m)}\right)^{2/3} \,, \tag{7}$$

$$R_{U_{k,m}} = R_{0_m}\sqrt{1 - \left(\frac{x_{k1}^{(m)}}{x_{U_m}}\right)^2} \,, \tag{8}$$

where $R_{0_m} = D_m + 8L_m$. Note that Eq. (8) implies an elliptical upwind dispersion. An overview of the main geometrical dimensions can be seen in Fig 4.

6) Downwind Concentration Ratio: B&M's nomogram provides a graphical way to obtain the downwind concentration ratio, $f_{k,m}$. The following procedure allows us to analytically approximate such value:

$$f_{k,m} = \begin{cases} f_{k,m}^{(1)}, & \phi_m \leq 0.2 \\ \left(\frac{5}{4} - \frac{5}{4}\phi_m\right) f_{k,m}^{(1)} + \left(\frac{5}{4}\phi_m - \frac{1}{4}\right) f_{k,m}^{(2)}, & 0.2 < \phi_m < 1 \\ f_{k,m}^{(2)}, & \phi_m \geq 1 \end{cases} \,, \tag{9}$$

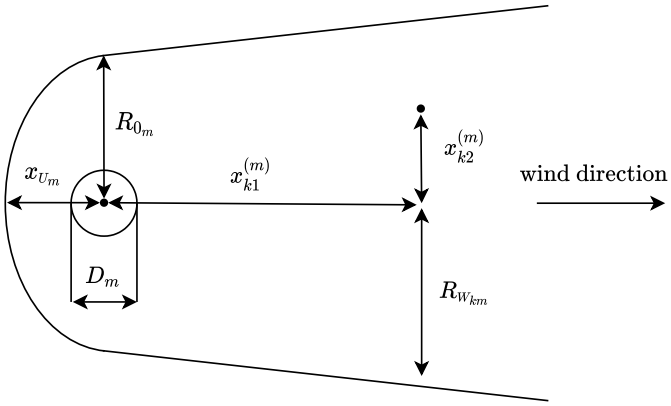


Fig. 4: Main geometrical dimensions used in B&M model.

where the functions $f_{k,m}^{(1)}$ and $f_{k,m}^{(2)}$ are defined as follows:

$$f_{k,m}^{(1)} \triangleq 306.25 \left(x_{k1}^{(m)}\right)^{-2} \dot{V}_m u^{-1} ,$$
$$f_{k,m}^{(2)} \triangleq 401.39 \left(x_{k1}^{(m)}\right)^{-1.95} \dot{V}_m^{0.78} {g'_m}^{-0.39} . \quad (10)$$

At low values of $x_{k1}^{(m)}$, Eq. (9) might result in $f_{k,m} > 1$, which is a wrong result generated by an excessive extrapolation from the original nomogram. Therefore, we provide the following correction to be applied to the result of Eq. (9):

$$\frac{f_{k,m}}{f_{k,m}+1} \longmapsto f_{k,m} . \quad (11)$$

7) Isobaric Transformation Correction: Finally, one last correction is necessary as the gas, once in contact with the atmosphere, begins a heat exchange (since $T_m < T_{atm}$) resulting in a temperature increase and expansion, therefore *decreasing* its concentration:

$$\frac{c_{k,m}}{c_{k,m} + (1 - c_{k,m})\frac{T_{atm}}{T_m}} \longmapsto c_{k,m} , \quad (12)$$

where T_{atm} and T_m are expressed as absolute temperatures. Eq. (12) implies an ideal gas behavior and a thermal equilibrium at any $x_{k1}^{(m)}$.

C. Signal Model

When a PRD opens, the released gas affects its value of concentration in the surrounding environment. The following equations describe the signal sensed by kth sensor in terms of concentration (y_k) in the case of normal operations (hypothesis $\mathcal{H}_0$), and in the case of an open PRD (hypothesis $\mathcal{H}_1$):

$$\begin{cases} \mathcal{H}_0 : & y_k = w_k \\ \mathcal{H}_1 : & y_k = c_{k,m} \cdot \xi_k + w_k \end{cases} , \quad (13)$$

where $w_k \sim \text{Gamma}\left(\frac{b^2}{\nu^2}, \frac{b}{\nu^2}\right)$ is the concentration present in the environment in normal conditions with b as its mean value and ν as its standard deviation. $c_{k,m}$ is the mean value of concentration contribution where the kth sensor is located due to the opening of the mth PRD. Finally, $\xi_k \sim \text{Gamma}\left(\omega^{-2}, \omega^{-2}\right)$ represents the fluctuation around the value of $c_{k,m}$, where ω is the relative mean fluctuation [22]. Due to the spatial separation of the sensors, it is assumed that ξ_k's and w_k's are both statistically independent. From Eq. (13), one can notice the following properties:

$$\mathbb{E}(y_k|\mathcal{H}_1; m) = c_{k,m} + b , \quad \mathbb{V}\text{ar}(y_k|\mathcal{H}_1; m) = \omega^2 c_{k,m}^2 + \nu^2 . \quad (14)$$

However, treating $p(y_k|\mathcal{H}_1; m)$ is not trivial, as it is the sum of two Gamma distributed random variables. Therefore, to simplify the the resulting algorithm, the sum of two independent Gamma random variables is approximated with a Gamma random variable with expectation (resp. variance) obtained as the sum of the expectations (resp. variances) of the two original variables (proof of validity in [23]). Consequently, during the development of the detectors, the distribution of the signal will be approximated with the following:

$$\begin{cases} y_k|\mathcal{H}_0 \sim \text{Gamma}\left(\alpha^{(0)}, \beta^{(0)}\right) \\ y_k|\mathcal{H}_1; m \overset{\text{approx.}}{\sim} \text{Gamma}\left(\alpha_{k,m}^{(1)}, \beta_{k,m}^{(1)}\right) \end{cases} , \quad (15)$$

where $\alpha^{(0)} \triangleq \frac{b^2}{\nu^2}$, $\beta^{(0)} \triangleq \frac{b}{\nu^2}$, $\alpha_{k,m}^{(1)} \triangleq \frac{(c_{k,m}+b)^2}{\omega^2 c_{k,m}^2 + \nu^2}$, and $\beta_{k,m}^{(1)} \triangleq \frac{c_{k,m}+b}{\omega^2 c_{k,m}^2 + \nu^2}$. Note that such approximated pdf becomes the actual pdf in case $c_{k,m} = 0$ when $\mathcal{H}_1$ is true.

III. Local Detection

As a consequence of Eq. (15), we can write the likelihoods of a sensor measurement:

$$p(y_k|\mathcal{H}_i) = \frac{{\beta^{(i)}}^{\alpha^{(i)}}}{\Gamma\left(\alpha^{(i)}\right)} y_k^{\alpha^{(i)}-1} e^{-\beta^{(i)} y_k} , \qquad i \in \{0, 1\} . \quad (16)$$

Note that $p(y_k|\mathcal{H}_1)$, $\alpha^{(1)}$, and $\beta^{(1)}$ are always referred to as $p(y_k|\mathcal{H}_1; m)$, $\alpha_{k,m}^{(1)}$, and $\beta_{k,m}^{(1)}$, respectively, to emphasize the dependency on the sensor and the source.

One should also keep in mind that Eq. (16), for $i = 1$, is an approximated pdf, and that such approximation will propagate throughout many of the equations in the rest of the work.

At a local detection level, the log-likelihood ratio test is *uniformly most powerful* (UMP) in a local sense having the following explicit expression for the test statistics:

$$\ln \frac{p(y_k|\mathcal{H}_1; m)}{p(y_k|\mathcal{H}_0)} = \left(\alpha_{k,m}^{(1)} \ln \beta_{k,m}^{(1)} - \alpha^{(0)} \ln \beta^{(0)}\right) - \left(\ln \Gamma\left(\alpha_{k,m}^{(1)}\right) - \ln \Gamma\left(\alpha^{(0)}\right)\right) + \left(\alpha_{k,m}^{(1)} - \alpha^{(0)}\right) \ln y_k - \left(\beta_{k,m}^{(1)} - \beta^{(0)}\right) y_k . \quad (17)$$

The last equation shows that an equivalent test is the following:

$$\ln y_k - \frac{\beta_{k,m}^{(1)} - \beta^{(0)}}{\alpha_{k,m}^{(1)} - \alpha^{(0)}} y_k \underset{d_k=0}{\overset{d_k=1}{\gtrless}} \gamma . \quad (18)$$

Eq. (18) shows that in order to perform a UMP test at local level, the detector needs to know the values of $\alpha_{k,m}^{(1)}$ and $\beta_{k,m}^{(1)}$. Unfortunately, these values are not available for the sensors as

they depend on the current wind speed and direction (which change over time), as well as the unknown parameter m. This leads to the conclusion that the test in Eq. (18) can be substituted with a local concentration test[5]:

$$y_k \underset{d_k=0}{\overset{d_k=1}{\gtrless}} \gamma \ . \tag{19}$$

Thanks to the approximation carried out in Eq. (15), it is easy to obtain the performance of the concentration test:

$$P_F \triangleq \Pr(y_k \geq \gamma | \mathcal{H}_0) = 1 - F_{y_k|\mathcal{H}_0}(\gamma) \ , \tag{20}$$

$$P_{D,k}(m) \triangleq \Pr(y_k \geq \gamma | \mathcal{H}_1; m) = 1 - F_{y_k|\mathcal{H}_1;m}(\gamma) \ . \tag{21}$$

The Neyman-Pearson approach is here employed to design the threshold γ by fixing the desired value of P_F, making the choice of the threshold independent of the the unknown parameter m and the considered sensor.

IV. Fusion Center Detection

A. Input's Likelihoods

According to the typology of input received by the FC (either $\boldsymbol{d}$ or $\boldsymbol{y}$), we can define two different expressions of the likelihoods at the FC.

1) Raw measurements as input: In such case we can combine Eq. (16) and the independence of the sensor measurements in space:

$$p(\boldsymbol{y}|\mathcal{H}_1; m) = \prod_{k=1}^{K} p(y_k|\mathcal{H}_1; m) \ . \tag{22}$$

Similarly, one can obtain $p(\boldsymbol{y}|\mathcal{H}_0)$ replacing $p(y_k|\mathcal{H}_1; m)$ with $p(y_k|\mathcal{H}_0)$.

2) Binary decisions as input: When the sensors transmit binary decisions it is clear that d_k follows a (conditional) Bernoulli distribution:

$$d_k|\mathcal{H}_0 \sim \mathcal{B}(P_F) \ , \qquad d_k|\mathcal{H}_1; m \sim \mathcal{B}(P_{D,k}(m)) \ . \tag{23}$$

In this case we can exploit the independence of the sensor decisions in space:

$$\Pr(\boldsymbol{d}|\mathcal{H}_1; m) = \prod_{k=1}^{K} \left[P_{D,k}(m)^{d_k} (1 - P_{D,k}(m))^{1-d_k} \right] \ . \tag{24}$$

Similarly, one obtains $\Pr(\boldsymbol{d}|\mathcal{H}_0)$ replacing $P_{D,k}(m)$ with P_F.

B. Centralized GLRT

In the case where the sensors transmit the raw measurements to the FC, a centralized GLRT (C-GLRT) fusion rule can be employed:

$$\begin{aligned} \Lambda_{\mathrm{C-GLRT}} &= \ln \frac{\max\limits_{m=1,\dots,M} p(\boldsymbol{y}|\mathcal{H}_1; m)}{p(\boldsymbol{y}|\mathcal{H}_0)} \\ &= \sum_{k=1}^{K} \left[\ln \frac{p(y_k|\mathcal{H}_1; \widehat{m}_{\mathrm{C-MLE}})}{p(y_k|\mathcal{H}_0)} \right] \underset{\widehat{\mathcal{H}}=\mathcal{H}_0}{\overset{\widehat{\mathcal{H}}=\mathcal{H}_1}{\gtrless}} \overline{\gamma} \ , \end{aligned} \tag{25}$$

[5]A local test based on the locally most powerful *score test* has been obtained, however computing $P_{D,k}(m)$ and P_F in closed-form is not possible.

where $\widehat{m}_{\mathrm{C-MLE}}$ is the Maximum Likelihood Estimate (MLE) among the possible source points conditioned to $\mathcal{H}_1$ being true:

$$\widehat{m}_{\mathrm{C-MLE}} = \underset{m=1,\dots,M}{\arg\max} \sum_{k=1}^{K} \ln p(y_k|\mathcal{H}_1; m) \ . \tag{26}$$

C. Distributed GLRT

When the sensors send a binary decision, the FC can perform a Generalized version of the well-known Chair-Varshney Rule, here named Distributed GLRT (D-GLRT):

$$\begin{aligned} \Lambda_{\mathrm{D-GLRT}} &= \ln \frac{\max\limits_{m=1,\dots,M} \Pr(\boldsymbol{d}|\mathcal{H}_1; m)}{\Pr(\boldsymbol{d}|\mathcal{H}_0)} \\ &= \sum_{k=1}^{K} \left[d_k \ln \frac{P_{D,k}(\widehat{m}_{\mathrm{D-MLE}})}{P_F} \right. \\ &\quad \left. + (1 - d_k) \ln \frac{1 - P_{D,k}(\widehat{m}_{\mathrm{D-MLE}})}{1 - P_F} \right] \underset{\widehat{\mathcal{H}}=\mathcal{H}_0}{\overset{\widehat{\mathcal{H}}=\mathcal{H}_1}{\gtrless}} \overline{\gamma} \ , \end{aligned} \tag{27}$$

where $\widehat{m}_{\mathrm{D-MLE}}$ is the MLE among the possible source points conditioned to $\mathcal{H}_1$ being true, namely:

$$\widehat{m}_{\mathrm{D-MLE}} = \underset{m=1,\dots,M}{\arg\max} \sum_{k=1}^{K} \ln \Pr(d_k|\mathcal{H}_1; m) \ . \tag{28}$$

D. Counting Rule

The well-known Counting Rule (CR) is among the simplest fusion rules, where the number of sensors detecting a dispersion is compared to a threshold:

$$\Lambda_{\mathrm{CR}} = \sum_{k=1}^{K} d_k \underset{\widehat{\mathcal{H}}=\mathcal{H}_0}{\overset{\widehat{\mathcal{H}}=\mathcal{H}_1}{\gtrless}} \overline{\gamma} \ . \tag{29}$$

V. Fusion Center Localization

A. Centralized/Distributed GLRT - Maximum Likelihood Estimator

In such a case, the source identification is *automatically* incorporated in the GLRT. Therefore, once obtained that $\widehat{\mathcal{H}} = \mathcal{H}_1$, the identified source will be $\widehat{m}_{\mathrm{C-MLE}}$ as defined in Eq. (26) for the centralized case. Analogously, for distributed GLRT, we can identify the source with $\widehat{m}_{\mathrm{D-MLE}}$ as defined in Eq. (28).

B. Counting Rule - Heuristic Estimators

Unlike the fusion rules based on the GLRT, the CR does not require the estimation of the parameter m. Therefore, a source identification method must be developed separately. A number of heuristic methods exist that can perform such task when the sole available information is $\boldsymbol{d}$ and $\{\boldsymbol{x}_k\}_{k=1,\dots,K}$. Here we investigate the use of the Centroid Method and the Center of the Minimum Enclosing Circle (CMEC) due to their popularity, simplicity, and effectiveness [2], [15], [24].

TABLE I: Computational Complexity of the Proposed Algorithms

Algorithm	Complexity
Local Concentration Test	$\mathcal{O}(1)$
C-GLRT / D-GLRT	$\mathcal{O}(KM)$
CR + Centroid / CMEC	$\mathcal{O}(K)+\mathcal{O}(K+M)$

1) Centroid Method: The Centroid Method calculates the centroid of the sensors detecting a release:

$$\boldsymbol{x}_{\mathrm{C}} = \frac{\sum_{k=1}^{K} d_k \boldsymbol{x}_k}{\sum_{k=1}^{K} d_k} \,. \tag{30}$$

However, the position $\boldsymbol{x}_{\mathrm{C}}$ may not correspond to any of the existing PRDs. Therefore, we infer that source point is the closest to the calculated position:

$$\widehat{m}_{\mathrm{CENTROID}} = \underset{m=1,\dots,M}{\arg\min} \|\boldsymbol{x}_C - \boldsymbol{h}_m\| \,. \tag{31}$$

2) CMEC Algorithm: The CMEC Algorithm calculates the center of the smallest circle enclosing all the sensors detecting a dispersion. Such point ($\boldsymbol{x}_{\mathrm{CMEC}}$) can be efficiently computed via *Megiddo Algorithm* (not reported here) [25]. Analogously to the Centroid Method, we need to employ the same final step to enforce the source position estimate to lie in the same discrete set:

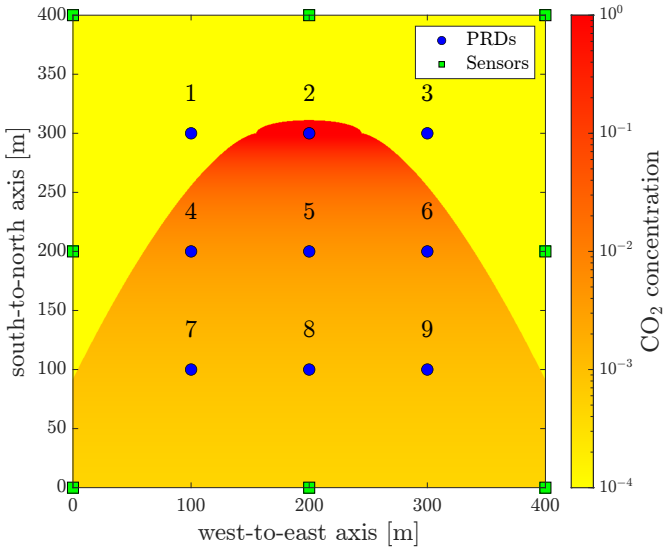
$$\widehat{m}_{\mathrm{CMEC}} = \underset{m=1,\dots,M}{\arg\min} \|\boldsymbol{x}_{\mathrm{CMEC}} - \boldsymbol{h}_m\| \,. \tag{32}$$

VI. Computational Complexity

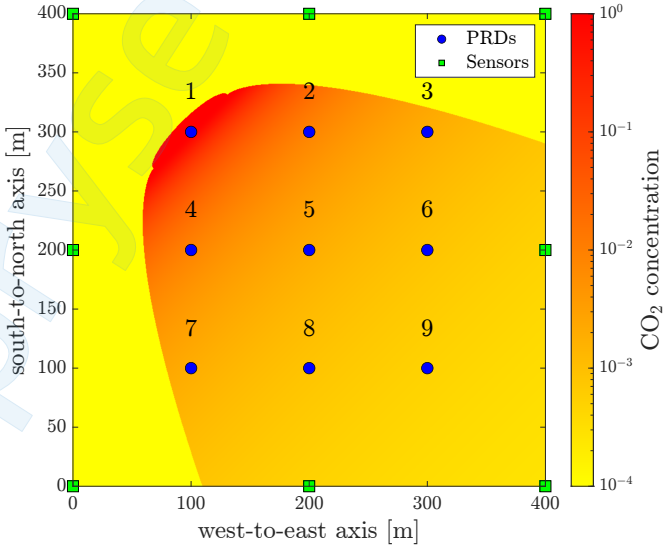
Tab. I shows the computational complexity of the proposed algorithms. It is noticeable how the local detection (when employed) is obtained with a finite number of operations. Moreover, we can notice that the detection methods based on the GLRT have higher complexity than the CR ($\mathcal{O}(KM)$ against $\mathcal{O}(K)$), which has the further advantage of performing the localization algorithms only when $\widehat{\mathcal{H}} = \mathcal{H}_1$. Note that both the Centroid Method and the CMEC Algorithm have the same complexity $\mathcal{O}(K+M)$. More specifically, obtaining $\boldsymbol{x}_{\mathrm{C}}$ and $\boldsymbol{x}_{\mathrm{CMEC}}$ (using *Megiddo Algorithm*) has complexity $\mathcal{O}(K)$, while both Eqs. (31) and (32) have complexity $\mathcal{O}(M)$.

VII. Simulation Results

The results are obtained simulating $M = 9$ identical sources placed in a square grid with side equal to 200 m. Such PRDs are located in a square-shaped plant of side $L = 400$ m monitored by $K = 8$ sensors installed on the perimeter and equally spaced. Such geometrical setup is illustrated in Fig. 5 while the rest of the simulation parameters are shown in Tab. II. Because of the symmetry properties of the chosen geometrical setup, only the wind directions in the interval $\varphi \in [\varphi_0, \varphi_0 + 45°]$ with $\varphi_0 \in [0°, 360°)$ can be evaluated, all other configurations can be mapped into one of those.



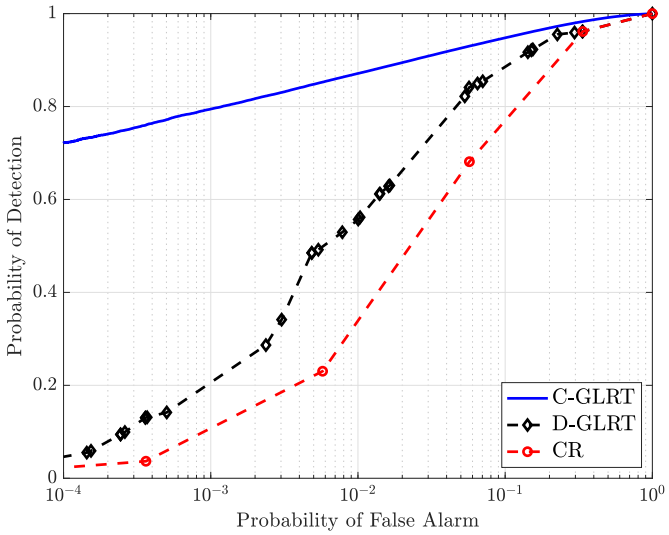
(a) Release from PRD ($m = 2$) and wind from north ($\varphi = 0°$).



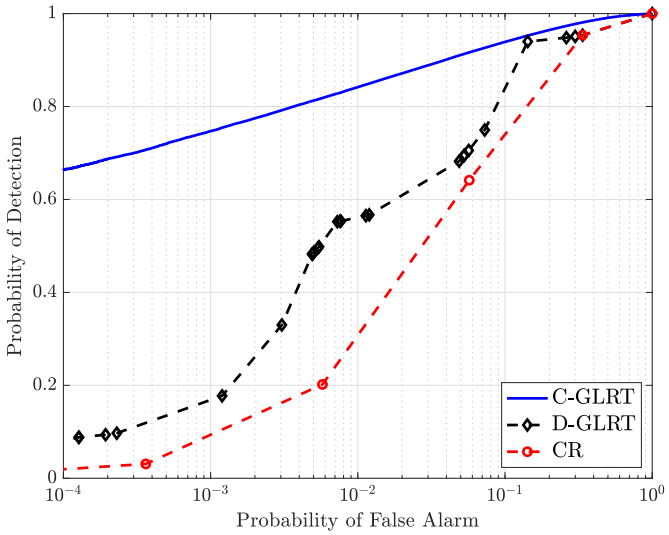
(b) Release from PRD ($m = 1$) and wind from north-west ($\varphi = 315°$).

Fig. 5: (Mean) concentration maps in a dispersion scenario.

For this reason we only considered $\varphi \in \{0°, 315°\}$. Because of the presence of $M = 9$ PRDs, the performances are averaged among all the release points, both in the detection and localization stage. From Figs. 6 and 7, it is immediately observable the superiority of the C-GLRT and its C-MLE estimator, as a consequence of directly transmitting $\boldsymbol{y}$ rather than $\boldsymbol{d}$ to the FC. However, a centralized network will likely show higher operating costs than a distributed network, especially in case of frequent measurements and transmissions to the FC. Therefore, the performance of the centralized network is solely used for benchmarking purposes. When considering one-bit quantization, we notice how the D-GLRT gives better performances than the CR in terms of detection (higher probability of detection given a fixed prob-

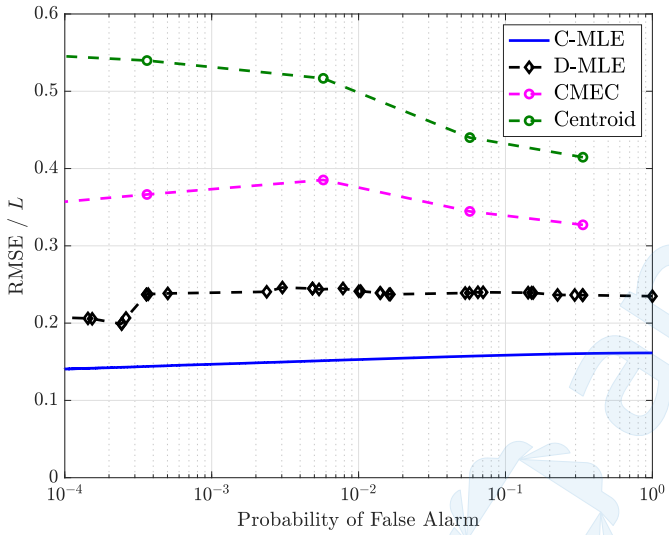


(a) Wind from north ($\varphi = 0°$).

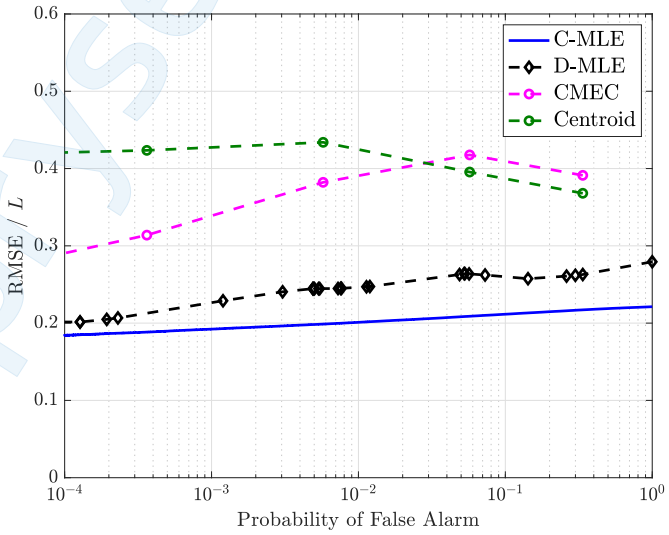


(b) Wind from north-west ($\varphi = 315°$).

Fig. 6: ROC curves at different wind directions.



(a) Wind from north ($\varphi = 0°$).



(b) Wind from north-west ($\varphi = 315°$).

Fig. 7: Localization performances at different wind directions.

ability of false alarm in the whole ROC space), highlighting how a model-aware design of the FC outperforms a heuristic design (since the CR can be used with no knowledge of the signal model). The localization performance follows a similar behavior showing how the D-MLE, in terms of *root mean square error* (RMSE), approaches the C-MLE, while the Centroid and the CMEC methods give worse results. In particular, the Centroid method gives typically the highest (i.e. worst) RMSE values. Nevertheless, an inversion of this trend can be seen in Fig. 7b: however, this holds only for values of probability of false-alarm higher than 0.057. From Fig. 7 it is evident how the 1-bit quantization makes it impossible to obtain a perfectly monotonic behavior of the RMSE as function of the probability of false alarm. On the contrary, the centralized configuration shows that lowering the detection threshold makes the system process less informative measurements, reducing the localization accuracy. While this tendency is somehow followed also by the D-MLE (some fluctuations of the RMSE are present among neighboring thresholds), a more unpredictable behavior is shown by the Centroid and CMEC methods. Ultimately, it is important to notice that while the CR allows only K thresholds, the D-GLRT allows $\left(2^K - 1\right) M$ thresholds, making it easier to tune the system to a desired false alarm rate[6], unless a randomization procedure is applied.

[6]The number of thresholds does not include those represented by the upper-right and lower-left corner points of the ROC curves.

TABLE II: Parameters used for the simulation

Parameter	Value	Note
c_m	1	$\forall m$, pure CO_2
$T_m^{(in)}$	253 K	$\forall m$, [11]
T_m	219 K	$\forall m$, *Soave-Redlich–Kwong* EOS [26]
T_{atm}	293 K	–
$P_m^{(in)}$	19.8 bar	$\forall m$
P_{atm}	1.0 bar	–
ρ_m	2.48 kg/m^3	$\forall m$, *Soave-Redlich–Kwong* EOS [26]
ρ_{air}	1.20 kg/m^3	[27]
u	1 m/s	–
$\dot{V}_m$	0.5312 m^3/s	$\forall m$
D_m	17.98 mm	$\forall m$
b	400 ppm	–
ν	200 ppm	–
ω	1	–
γ	985 ppm	from Eq. (20) with $P_F = 0.05$

VIII. Conclusions and Future Directions

In this work, we addressed the distributed detection (via a WSN) of CO_2 release from storage tanks caused by the opening of PRDs and localization of the corresponding activated PRD. The sensors individually monitor the facility and transmit their decisions to a FC based on an individual concentration level test. Herein, a spatial aggregation is carried out based on GLRT and a global decision is performed. Results have highlighted the benefit in terms of ROC with respect to the well-known CR that does not include the knowledge of the dispersion model in its design. Similar benefits have been observed for MLE-based estimators as compared to naive alternatives based on CMEC and centroid approaches. Future directions will include: (i) sequential/quickest detection setups and corresponding localization techniques, (ii) the combined adoption of quantization with censoring techniques [28], (iii) more complex dispersion models, and (iv) the use of channel-aware techniques [29].

References

[1] L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Trans. Industr. Inform.*, vol. 10, no. 4, pp. 2233–2243, 2014.

[2] G. Tabella, N. Paltrinieri, V. Cozzani, and P. Salvo Rossi, "Wireless sensor networks for detection and localization of subsea oil leakages," *IEEE Sens. J.*, vol. 21, no. 9, pp. 10 890–10 904, 2021.

[3] M. Ortner and A. Nehorai, "A sequential detector for biochemical release in realistic environments," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4173–4182, 2007.

[4] G. Tabella, D. Ciuonzo, N. Paltrinieri, and P. Salvo Rossi, "Spatio-temporal decision fusion for quickest fault detection within industrial plants: The oil and gas scenario," in *IEEE 24th Int. Conf. Inf. Fusion (FUSION)*, 2021.

[5] T. Zhao and A. Nehorai, "Detecting and estimating biochemical dispersion of a moving source in a semi-infinite medium," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2213–2225, 2006.

[6] S. Aldalahmeh, M. Ghogho, and A. Swami, "Fast distributed detection, localization, and estimation of a diffusive target in wireless sensor networks," in *7th IEEE Int. Symp. Wirel. Commun. Syst. (ISWCS)*, 2010, pp. 882–886.

[7] B. Ristic, A. Gunatilaka, and R. Gailis, "Achievable accuracy in Gaussian plume parameter estimation using a network of binary sensors," *Inf. Fusion*, vol. 25, pp. 42–48, 2015.

[8] S. Vijayakumaran, Y. Levinbook, and T. F. Wong, "Maximum likelihood localization of a diffusive point source using binary observations," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 665–676, 2007.

[9] Y. Di Martino, S. E. Duque, G. Reniers, and V. Cozzani, "Making the chemical and process industries more sustainable: Innovative decision-making framework to incorporate technological and non-technological inherently safer design (ISD) opportunities," *J. Clean. Prod.*, vol. 296, p. 126421, 2021.

[10] P. Harper, *Assessment of the major hazard potential of carbon dioxide (CO_2)*. UK: Health and Safety Executive, 2011.

[11] NIST, "Thermophysical Properties of Carbon dioxide," 2022.

[12] B. Ristic, A. Gunatilaka, and R. Gailis, "Localisation of a source of hazardous substance dispersion using binary measurements," *Atmos. Environ.*, vol. 142, pp. 114–119, 2016.

[13] D. D. Selvaratnam, I. Shames, D. V. Dimarogonas, J. H. Manton, and B. Ristic, "Co-operative estimation for source localisation using binary sensors," in *56th IEEE Annu. Conf Decis Control (CDC)*, 2017, pp. 1572–1577.

[14] 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.

[15] A. Shoari, G. Mateos, and A. Seyedi, "Analysis of target localization with ideal binary detectors via likelihood function smoothing," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 737–741, 2016.

[16] R. E. Britter and J. McQuaid, *Workbook on the dispersion of dense gases*. UK: Health and Safety Executive, 1988.

[17] S. Hanna, "Britter and McQuaid (B&M) 1988 workbook nomograms for dense gas modeling applied to the Jack Rabbit II chlorine release trials," *Atmos. Environ.*, vol. 232, p. 117539, 2020.

[18] TNO, *Yellow Book – Methods for the calculation of physical effects*. The Hague, The Netherlands: The Committee for the Prevention of Disasters by Hazardous Materials, 2005.

[19] S. Mannan, *Lees' Loss Prevention in the Process Industries*, 4th ed. Oxford, UK: Butterworth-Heinemann, 2012.

[20] CCPS, *Guidelines for Consequence Analysis of Chemical Releases*, 1st ed. Hoboken, NJ, USA: John Wiley & Sons, 1999.

[21] D. A. Crowl and J. F. Louvar, *Chemical Process Safety: Fundamentals with Applications*, 4th ed. London, UK: Pearson Education, 2019.

[22] M. Cassiani, M. B. Bertagni, M. Marro, and P. Salizzoni, "Concentration fluctuations from localized atmospheric releases," *Bound.-Layer Meteorol.*, vol. 177, no. 2, pp. 461–510, 2020.

[23] F. Babich and G. Lombardi, "Statistical analysis and characterization of the indoor propagation channel," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 455–464, 2000.

[24] A. Shoari and A. Seyedi, "Localization of an uncooperative target with binary observations," in *11th IEEE Int. Workshop Signal Process. Adv. Wirel. Commun. (SPAWC)*, 2010.

[25] N. Megiddo, "Linear-time algorithms for linear programming in r3 and related problems," in *23rd Annu. Symp. Found. Comput. Sci. (SFCS 1982)*, 1982, pp. 329–338.

[26] J. M. Smith, H. C. Van Ness, M. M. Abbott, and M. T. Swihart, *Introduction to Chemical Engineering Thermodynamics*, 8th ed. New York (NY), USA: McGraw-Hill Education, 2018.

[27] Engineering ToolBox, "Air - Density, Specific Weight and Thermal Expansion Coefficient vs. Temperature and Pressure," 2022.

[28] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr 1996.

[29] M. A. Al-Jarrah, M. A. Yaseen, A. Al-Dweik, O. A. Dobre, and E. Alsusa, "Decision fusion for IoT-based wireless sensor networks," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1313–1326, 2019.