*Article*

# Distance Based Analysis of Early Fire Indicators on a New Indoor Laboratory Dataset with Distributed Multi-Sensor Nodes

**Pascal Vorwerk [1,*], Jörg Kelleter [2], Stefen Müller [2] and Ulrich Krause [1]**

[1] Faculty of Process- and Systems Engineering, Institute of Apparatus and Environmental Technology, Otto von Guericke University of Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany
[2] GTE Industrieelektronik GmbH, Helmholtzstr. 21, 38-40, 41747 Viersen, Germany
* Correspondence: pascal.vorwerk@ovgu.de

**Abstract:** This work analyzes a new indoor laboratory data set looking at early fire indicators in controlled and realistic experiments representing different incipient fire scenarios. The experiments were performed within the confines of an indoor laboratory setting using multiple distributed sensor nodes on different room positions. Each sensor node collected data of particulate matter (PM), volatile organic compounds (VOC), CO, $CO_2$, $H_2$, UV, air temperature and humidity in terms of a multivariate time series. These data hold immense value for researchers within the machine learning and data science communities who are keen on exploring innovative advanced statistical and machine learning techniques. They serve as a valuable resource for the development of early fire detection systems. The analysis of the collected data was carried out in dependence of the manhattan distance between the fire source and the sensor node. We found that especially larger particles (> 0.5 $\mu$m) and VOCs show a significant dependency with respect to the intensity as a function of the manhattan distance to the source. Moreover, we observed differences in the propagation behavior of VOCs, PM, and CO, which are particularly relevant in incipient fire scenarios due to the presence of strand popagation effects.

**Keywords:** early fire detection; multi sensor network; data driven fire detection; machine learning; public data set

## 1. Introduction

The increasing availability of (low-cost) sensor technology has resulted in high amounts of data available for numerous processes. [1–4] This data can be accessed in near real-time, and advancements in computing technology, particularly in the realm of smart computing, enable the processing of large volumes of data to extract crucial information. I Consequently, it becomes possible to continuously monitor complex systems such as the occurrence of fires in buildings. [5–10] Early fire detection, in particular, has garnered significant interest, with multi-sensor approaches showing promising potential when compared to traditional smoke alarm devices. [11]

In recent decades, the interest in developing machine learning-based fire detection solutions has grown significantly to handle the expanding volume of multi-sensor data. [5,12] However, there is still a scarcity of available real fire data sets for model building and validation [11,13,14] due to the high associated costs. One drawback of the existing real fire data sets is the lack of diverse distances between sensors and the location of the fire source since multi-sensor fire detection systems are constrained to a specific detection area [6].

Previous studies have predominantly employed a single sensor placed on the ceiling of a fire test room, coupled with a fixed position for the fire source [11,14–16]. In some other studies, the propagation effects were limited by using small room sizes [6,17]. Gutmacher et al. [18] investigated the differences in the propagation behavior of smoke particles and gases in various EN54 standard test fires. They utilized a multi-sensor array and commercial smoke detectors to measure CO, NO2,

 doi:10.20944/preprints202306.0588.v1



humidity, temperature, and smoke density. The authors noted a paucity of research on examining the varying propagation behavior of gases and smoke during (incipient) fire scenarios.

Considering the inherent difficulty of accurately determining the location of a fire source [19,20] and the resulting distance between the fire origin and the sensor in real-world fire detection applications, it is crucial for models to be robust in the face of such uncertainty.

To fully harness the potential of multi-sensor measurements, encompassing gases, vapors such as volatile organic components (VOC), and particulate matter (PM), as early fire indicators, it is imperative to develop a comprehensive understanding of how these sensor measurements interact in various incipient fire scenarios and at different distances from the fire source. [21]

In this work, we provide and analyse a new data set that allows model building and validation at various sensor node positions in a standard EN54 test room using a multi-sensor node network. For the first time, this enables the consideration of strand propagation effects during the initial stages of various incipient fires within a real fire data set.

The subsequent sections of this paper introduce the experimental setup utilized in this study. Following that, we examine the behavior of sensor measurements including PM, VOC, CO, $CO_2$, $H_2$, UV, air temperature, and humidity. Specifically, we focus on their dispersion characteristics and the time it takes for them to become relevant as early fire indicators in various incipient fire scenarios. In conclusion, we assess the usability of the investigated early fire indicators in terms of their robustness to position dependency and their contribution to achieving the earliest possible detection.

## 2. Related Work

In the previous literature, a variety of different sensor measurements was used in terms of "fire parameters" [22] for (early) fire detection. In the context of this paper, we propose using the term "early fire indicators" in sense of direct sensor measurements derived from a multi sensor approach for detecting incipient fires at early stages. Additional early fire indicators that are model-based (e.g. trend features) are part of an additional work.

Previous study's often used a combination of carbon monoxide (CO) and carbon dioxide ($CO_2$) in the sense of early fire indicators. [11,15,17]

Other works proposed using air temperature, humidity and smoke concentration in addition to CO and $CO_2$. [5,16]

Wu et al. [5] argued that carbon dioxide ($CO_2$) and humidity are relevant "byproducts" of combustion processes but excluded these measurements from there work since they are highly affected by environmental changes and therefore prone to false positives.

Solórzano et al. [11] implemented hydrogen ($H_2$) as well as methane ($CH_4$), nitrogen oxides ($NO_x$) and VOC´s as additional early fire indicators. The authors found that early emissions of incipient fires (smoldering fires) include high levels of CO and VOC´s. The authors stated the importance of measuring VOC within a multi-sensor approach for detecting fires in a very early stage. [11]

This founding was confirmed by Nazir et al. [14] who used total volatile organic compounds (TVOC) in addition to air temperature, humidity, CO2 and ammonia (NH3) as early fire indicators.

As Rachman et al. [6] stated, multi- sensor fire detection systems are limited to the detection area. The authors proposed using multiple sensor nodes within one room to improve the detection area and the sensor sensitivity for the application room. The authors placed a total of 8 sensor nodes in every corner of a rectangular test room. Each sensor node contained a "[...] fire sensor, smoke sensor, and temperature sensor" [6] which were not specified in more detail. However, the size of the test room was very small – (50 x 50 x 60) $cm^3$ – which is why it is questionable whether the authors were able to detect relevant differences in the propagation behavior of the sensor measurements.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 June 2023 doi:10.20944/preprints202306.0588.v1



## 3. Materials and Methods

### 3.1. Sensor Network Setup

The sensor network setup used within the experimental studies can be seen in Figure 1. The sensors of each sensor node are controlled by a microcontroller (ESP32). The microcontroller communicates via WiFi and the protocol MQTT with a broker/server. The server was physically a Raspberry Pi.

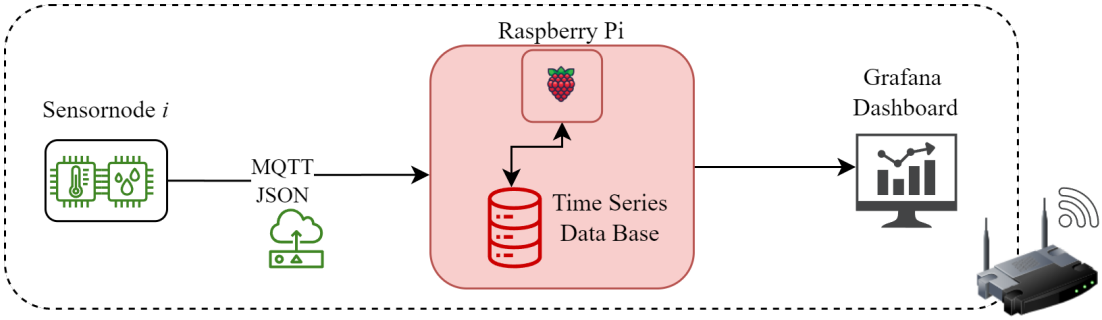


**Figure 1.** Sensor network architecture for the experimental setup.

The microcontroller sent the sensor data in the form of a JSON to the Raspberry Pi where a Python script decoded the JSON and wrote the sensor data into an Influx time series database where an UTC timestamp is automatically assigned to each measuring point. The sampling time was 1 sample per 10 seconds. Measurement data were exported from the Influx time series database as CSV-file using a Python script. For real time monitoring of the sensor data during the experiments, a Grafana dashboard was used. Each node in the network contained the sensors listed in Table 1.

**Table 1.** Overview of sensors in each sensor node.

| Sensor | Manufacture | Measurand | Unit |
|---|---|---|---|
| SPS30 | Sensirion | Particulate Matter | $cm^{-3}$ |
| SVM40 | Sensirion | VOC (Volatile organic compounds) | A.U. |
| CO/MF-1000 | MEMBRAPOR | $CO$ | $ppm$ |
| UST6xxx | UST | $H_2$ | $ppm$ |
| SCD40 | Sensirion | $CO_2$ | $ppm$ |
| UVTRON | HAMAMATSU | UV photon | # |
| SGP40 | Sensirion | Temperature, Relative air humidity | °C, % |

According to Rachman et al. [6], we employed a sensor network using various distributed sensor nodes in our experimental setting. The network consisted of 9 sensor nodes (0008 - 0016) that were positioned around the fire source, as shown in Figure 2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 June 2023 doi:10.20944/preprints202306.0588.v1



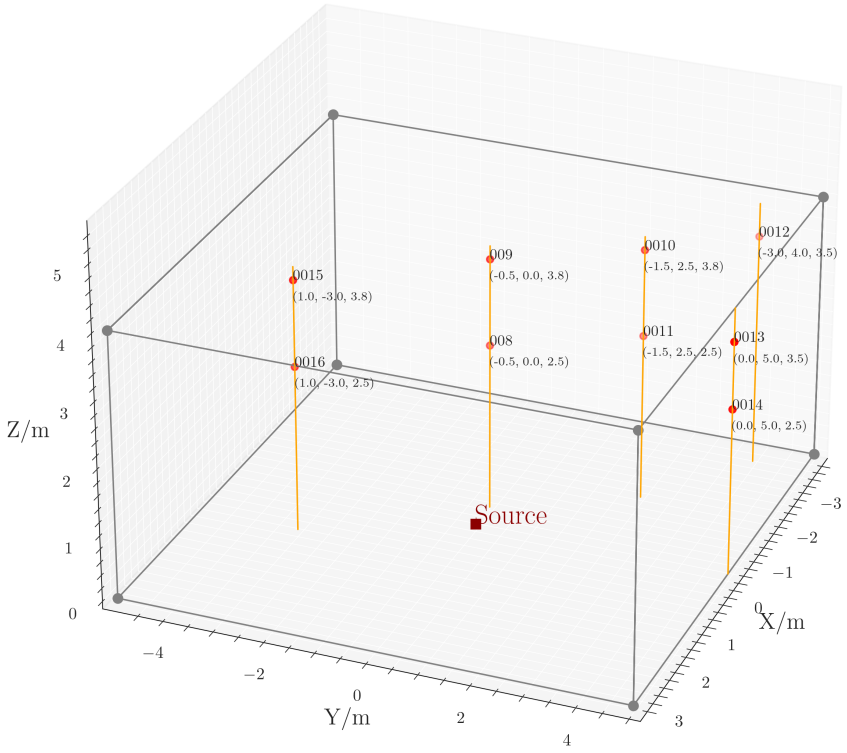


**Figure 2.** Positions of the fire source and sensor nodes in the EN54 fire test room.

All experiments were conducted in a non-ventilated EN54 fire test room with dimensions of (7 x 10 x 4) $m^3$. The positions of the sensor nodes in the fire room resulted in the following distances (Euclidean and Manhattan) with respect to the fire source (see Table 2).

**Table 2.** Positions and distances of sensor nodes and the source within the test room.

| Sensor_ID | x | y | h | euclidean | manhattan |
|---|---|---|---|---|---|
| | [m] | [m] | [m] | [m] | [m] |
| SK_8 | -0.5 | 0.0 | 2.5 | 2.5 | 3.0 |
| SK_9 | -0.5 | 0.0 | 3.8 | 3.8 | 4.3 |
| SK_10 | -1.5 | 2.5 | 3.8 | 4.8 | 7.8 |
| SK_11 | -1.5 | 2.5 | 2.5 | 3.8 | 6.5 |
| SK_12 | -3.0 | 4.0 | 3.5 | 6.1 | 10.5 |
| SK_13 | 0.0 | 5.0 | 3.5 | 6.1 | 8.5 |
| SK_14 | 0.0 | 5.0 | 2.5 | 5.6 | 7.5 |
| SK_15 | 1.0 | -3.0 | 3.8 | 4.9 | 7.8 |
| SK_16 | 1.0 | -3.0 | 2.5 | 4.0 | 6.5 |

To ensure a more accurate assessment of the plume-shaped propagation behavior, we adopt the Manhattan distance for further evaluation instead of the Euclidean distance. This distance measure takes into account the varying x-y distances between the sensor nodes 009 (located above the source) and 0011 (or 0016) and the source itself. While the Euclidean distance may yield similar values for these nodes, the Manhattan distance provides a more precise representation of the propagation behavior of the released products by accurately capturing the varying x-y distances to the source.

### *3.2. Experimental Procedure*

The experiments carried out can be divided into fire scenarios, nuisance scenarios and background measurements, as summarized in Table 3

 doi:10.20944/preprints202306.0588.v1



**Table 3.** Overview of the experiments carried out looking at the termination criterion and the number of experiments.

| Scenario | Termination Criterion | Number of Experiments |
|---|---|---|
| Wood | Max. Duration of Experiment | 3 |
| Candles | Max. Duration of Experiment | 3 |
| Cable | Max. Duration of Experiment | 3 |
| Lunts | Max. Duration of Experiment | 3 |
| Ethanol | Sample completely evaporated | 3 |
| Deodorant | Two Sprays of 15 Seconds | 2 |
| Hairspray | Two Sprays of 15 Seconds | 1 |
| Background | - | - |
| | | Total: 18 |

The experimental series consisted of 4 fire experiments as shown in Figure 3, 3 nuisance experiments and multiple background measurements between fire and nuisance experiments. The fire experiments were carried out as follows.

(a) (b) (c) (d)

**Figure 3.** Setup of carried out fire experiments; (a) Wood, (b) Candles, (c) Cable, (d) Lunts.

**Wood fire** was performed with a beech wood with dimensions (6 x 8 x 0.9) cm and an quartz radiant heater ("HQH 1200/3/1") from the company "ROWI" with a heating capacity of 1200 W. The quartz radiant heater was positioned at a distance of 5 cm from the vertical beech wood. With this test arrangement a wood smoldering fire was simulated. A flaming combustion was not triggered. The average wood moisture content was approx. 5 %. The prepared beech wood pieces where stored dry at standard conditions prior to the experiments. The wood samples had a mean mass of 30.8 g before starting the experiment procedure.

**Candle fire** was performed with four commercially available tea lights (100 % paraffin, housing included) which where placed on a heating plate. The mean mass of the four candles was 47.63 g. To accelerate the melting process, the heating plate was heated to a constant temperature of 450 °C approx. 30 s after lighting the tea lights.

**Cable fire** was performed with an halogen-free NHXMH-type cable (length = 35 cm) from manufacture "Lapp Group". The fire was triggered by an electrical overload of approx. 130 A DC. The cable was insulated on the lower side with rock wool.

**Lunts fire** was performed with eight lunts of cutton material (diameter = 4 mm) from manufacture "Fehr-Erlen". The average mass of the 8 fuses before the start of the experiment was 200.40 g.The Lunts where arranged annularly and ignited by a heating coil with a surface temperature of approx. 600 °C.

After fire experiments, nuisance measurements were performed to create a set of non-fire data.

**Ethanol** was evaporated from a vat above the heating plate. For this purpose, the hot plate was heated to 50 °C. A total volume of 10 ml ethanol per experiment was completely evaporated.

**Deodorant** was applied evenly into the room by means of two 15 s sprays. We used deodorant from manufacture "elkos" without aluminum salts.

 doi:10.20944/preprints202306.0588.v1



**Hairspray** was applied equivalent to the deodorant with two sprays of 15 s. We used hairspray from manufacture "Schwarzkopf", type "3 Wetter taft ULTRA 4".

**Background** was measured between nuisance/ fire experiments for a minimum measurement time of one hour. The background measurement started when ventilating of the test room has been done and a minimum waiting time of 30 minutes has been passed.

### *3.3. Data Set*

The data set contains 305304 rows and 16 columns in total and is structured as a continuous multivariate time series. Each row contains all measurements from one unique sensor node at a unique time stamp. The columns represent the sensor measurements explained in Table A1.

The data is labeled with four types of labels: a scenario specific label ("scenario_label"), a binary label ("anomaly_label") which distinguishes between a "Normal" state (background) and an "Anomaly" (fire or nuisance scenario), a ternary label ("ternary_label") distinguishing between "Nuisance", "Fire" and "Background" as well as a progress label ("progress_label") that can be used to divide the sequences of events into sub-sequences depending on the ongoing physical sub-processes.

The data set contains 82.98 % background data points and 17.02 % anomaly data points which can be divided into 12.50 % data points of fire anomalies and 4.52 % data points of nuisance anomalies.

## 4. Results

The response behavior of the sensor measurements was evaluated based on two criteria. The first criterion involved determining the maximum absolute value of each measurement, considering the scenario and the Manhattan distance to the source (as discussed in Section 4.1). The second criterion involved analyzing the time difference between the start of the experiment and the point at which a defined threshold was reached, taking into account the scenario and the Manhattan distance to the source (as described in Section 4.2).

### *4.1. Intensity Dependence*

Figure 4 depicts the maximum values obtained for each sensor measurement, considering both the scenario and the Manhattan distance from the source. In general, a decrease in intensity with increasing Manhattan distance between sensor node and the source is visible.

However, there are some notable exceptions observed for the $CO_2$, humidity, UV, and PM05 measurements. In the case of the conducted fire experiments, which primarily involved incomplete combustion, only small amounts of $CO_2$ were released, resulting in minimal impact on the $CO_2$ measurement.

An exception to this trend is observed in the candles scenario, where more complete combustion takes place, leading to higher levels of $CO_2$ release. Interestingly, in the candles scenario, the dispersion of $CO_2$ appears to be nearly independent of the distance between the source and the sensor. On the other hand, humidity remained largely unaffected by the conducted experiments, exhibiting no discernible trend in relation to the distance from the source.

Analyzing the UV measurements, two noteworthy conclusions can be drawn. Firstly, UV signals were exclusively detected in the Candles scenario, which was the only scenario characterized by flaming combustion. This indicates that UV measurements can serve as a reliable indicator for scenarios involving open flames. Secondly, the maximum values of the UV measurements were unaffected by the Manhattan distance from the source. This can be attributed to the absence of any objects within the room that would obstruct the "line of sight" between the sensor node and the source, allowing the UV signals to propagate freely regardless of the distance.

The air temperature exhibits a minor decline as the distance from the source increases, albeit within a narrow range of 24°C to 28°C. This observation can be attributed to the nature of the fire scenarios conducted, which primarily involved fires in their initial stages. As a result, the heat release rates were generally low, leading to only a marginal impact on the surrounding air temperature.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 June 2023 doi:10.20944/preprints202306.0588.v1



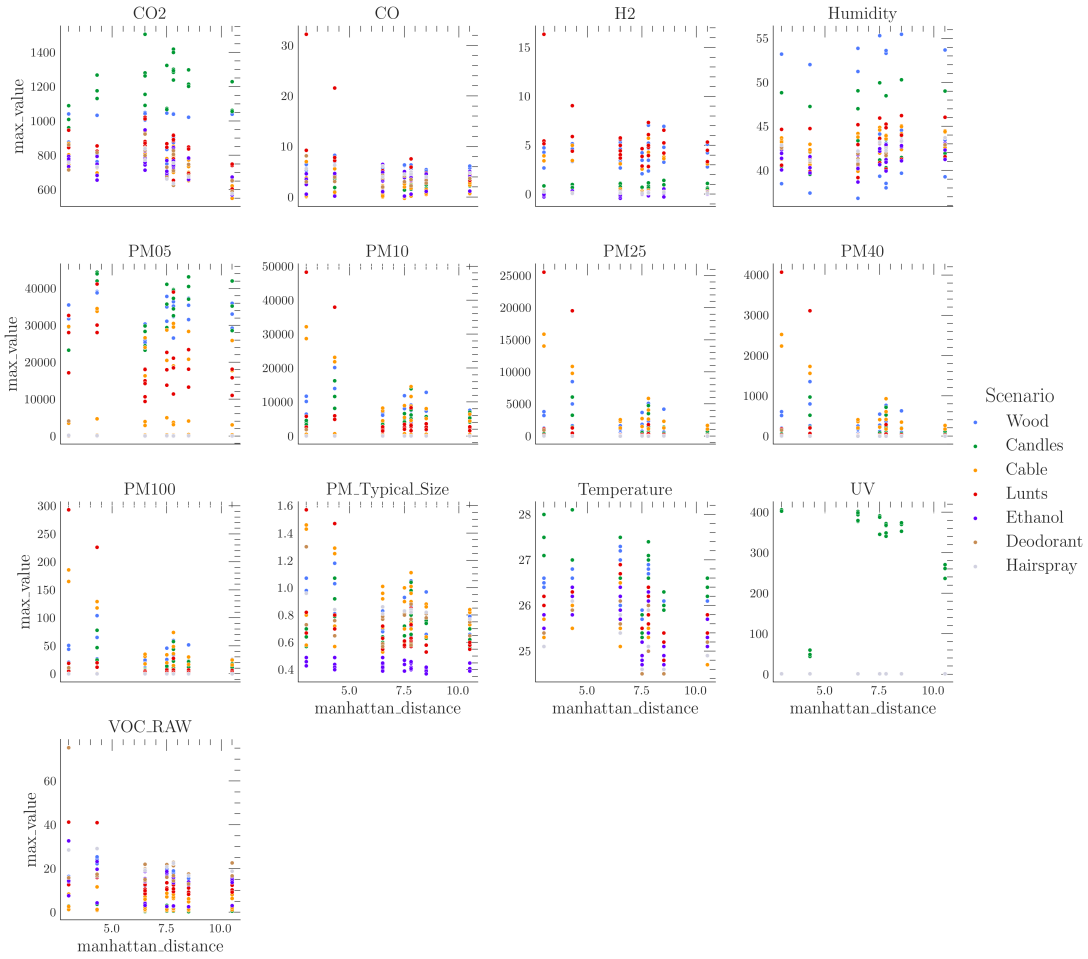


**Figure 4.** Absolute maximum values for each sensor measurement in dependence on the Manhattan distance to the source and the scenario.

VOC exhibit a noticeable correlation with the Manhattan distance from the source. Figure 4 demonstrates that VOC is released in the lunts, cable, and wood scenarios, as well as in all nuisance scenarios involving ethanol, deodorant, and hairspray. However, VOC is not detected in the candles scenario.

Remarkably, PM exhibited distinct behavior: smaller particles (PM05) showed minimal sensitivity to the distance from the source, while the count of larger particles (PM10-PM100) decreased as the distance increased. Two primary factors contribute to this pattern: agglomeration and gravitational settling, as discussed by Mostafa et al. [23].

Agglomeration refers to the process by which particles stick together to form larger clumps of particles. Agglomeration can result in the formation of particles that are too large to remain suspended in the air, leading to a decrease in the concentration of PM. [24]

The extent to which gravitational settling affects particles depends on their size and density. According to Mostafa et al. [23], larger particles exhibit higher sedimentation velocities compared to smaller particles, and the impact of gravity on smaller particles is relatively smaller in terms of their propagation behavior.From this, we can conclude that smaller particles are more susceptible to air movement,and their behavior is not primarily affected by agglomeration, as it would result in the opposite observation (a decrease in the number of smaller particles and a less pronounced decrease in the number of larger particles with increasing Manhattan distance from the source).

In our study, the effect of gravitational settling appears to be the primary determinant of particle propagation, as there were no external influences such as forced ventilation or significant temperature

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 June 2023 doi:10.20944/preprints202306.0588.v1



variations within the test room. The larger particles tend to move closer to the ground due to gravity, leading to less frequent detection by the more distant sensor nodes.

Furthermore, we made an intriguing observation regarding sensor nodes positioned at the same height (2.5 m) and with an equal Manhattan distance from the source (e.g., sensor node 0011 and sensor node 0016, both at a Manhattan Distance of 6.5 m), but located at different positions in the x-y direction of the room. Within a given fire scenario, these nodes exhibited slightly varied maximum values for each measurement. A specific example can be seen in the wood scenario illustrated in Figure 5. It is evident that sensor node 0011 recorded slightly higher maximum values for CO and all PM (excluding PM05), while displaying lower maximum values for VOC compared to sensor node 0016.

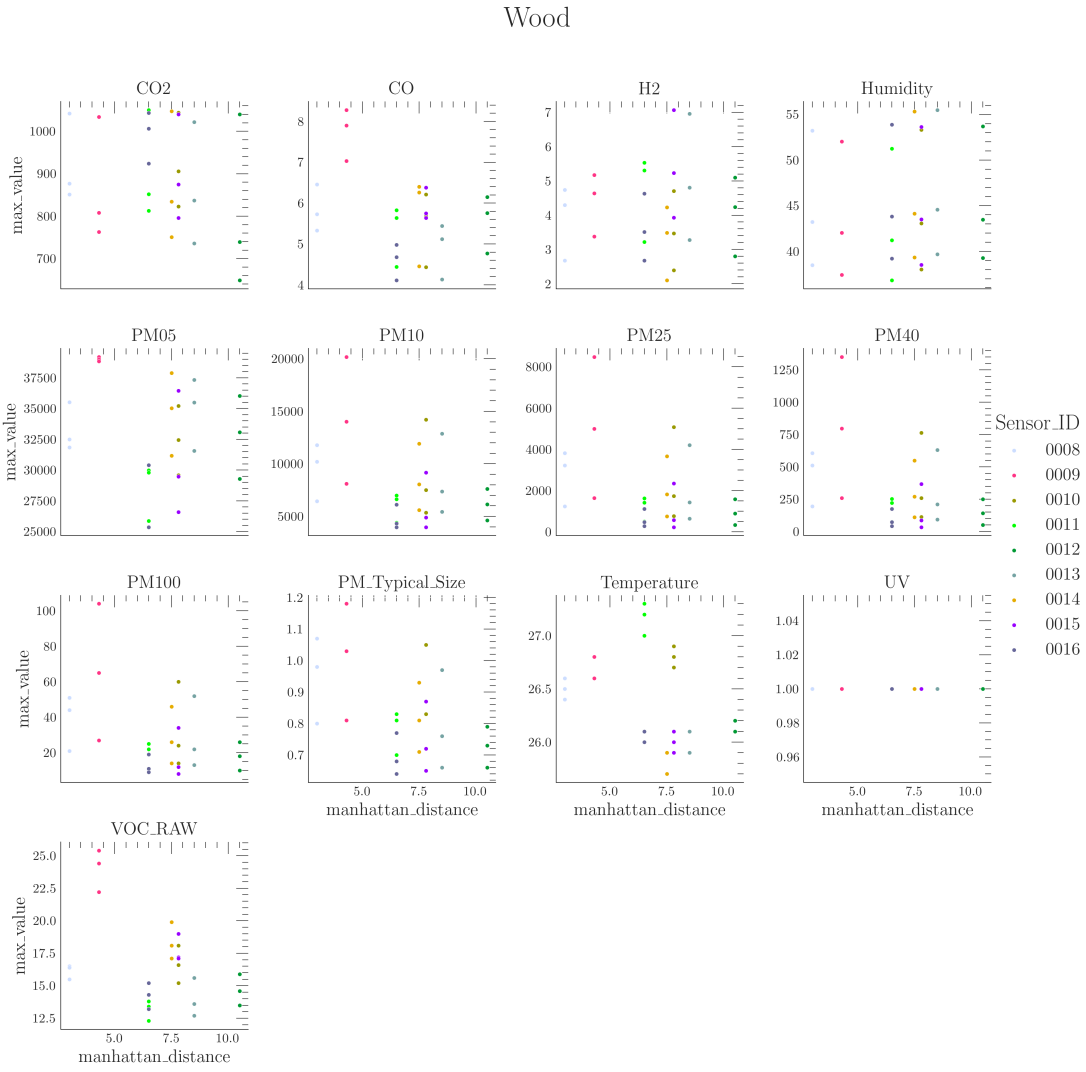


**Figure 5.** Absolute maximum values for each sensor measurement in dependence of the Manhattan distance to the source and the Sensor_ID within the wood scenario.

When comparing sensor nodes 0015 and 0010 (Manhattan distance = 7.8 m, height = 3.8 m) positioned directly above sensor nodes 0016 and 0011, we notice slight variations in their behavior. Sensor node 0010 recorded higher values of PM (excluding PM05) and VOCs compared to sensor node 0015. This aligns with the behavior of sensor nodes 0016 and 0011, located below, except for the PM05 measurement. Figure 4 indicates that PM05 does not exhibit a clear dependence on the distance from the source in general, suggesting that the exception here may be due to random variation in propagation behavior.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 June 2023 doi:10.20944/preprints202306.0588.v1



However, when it comes to CO measurement, sensor node 0010 registered slightly lower maximum values compared to sensor node 0015, which contrasts with the behavior of sensor nodes 0016 and 0011 (where sensor node 0011 recorded higher maximum CO values than sensor node 0016). Given the absence of external ventilation effects and consistent temperature conditions within the test room, we can infer that different factors are influencing the propagation behavior of combustion products. PM and VOCs seem to propagate uniformly, but CO behaves differently.

One possible explanation is that the VOCs released during combustion processes primarily consist of vapor components, as stated in [25] and [26]. These components, similar to PM, are heavier than pure gases and likely disperse differently compared to the released gaseous components. It is plausible that multiple dispersion patterns occur, with some strands primarily containing gases, while others include VOCs and larger particles.

The central assumption is that VOC´s, PM, and gases (such as CO) form distinct strands during the combustion process, potentially at different stages. It should be noted that these strands are not comprised solely of a single component, but rather have a predominant presence of a component.

### *4.2. Time Dependence*

In order to examine the time delay associated with different sensor node positions, a global threshold was established for each sensor measurement. The threshold values for detection time are outlined in Table 4.

**Table 4.** Global thresholds for every sensor measurement.

| Measurand | Global Threshold | Unit |
|---|---|---|
| $CO_2$ | 1200 | ppm |
| CO | 2 | ppm |
| $H_2$ | 1.5 | ppm |
| Humidity | 75 | % |
| PM05 | 100 | $cm^{-3}$ |
| PM10 | 75 | $cm^{-3}$ |
| PM25 | 75 | $cm^{-3}$ |
| PM40 | 75 | $cm^{-3}$ |
| PM100 | 10 | $cm^{-3}$ |
| PM_Typical_Size | 0.55 | $\mu$m |
| Temperature | 30 | °C |
| UV | 3 | # |
| VOC_RAW | 2.5 | A.U. |

The determination of the thresholds involved prior analysis of the time series data collected. The aim was to establish values that would only be surpassed by exceptional peaks associated with specific experiments, ensuring robustness. At the same time, the thresholds were set as low as possible to enable detection even at sensor nodes located far from the emission source, thereby enhancing sensitivity.

The thresholds were established using a rule-based baseline model designed for anomaly detection. They were set globally, irrespective of the scenario or the position of the sensor node. This decision was made because in real-world applications, the specific scenario and the distance between the sensor node and the emission source are typically unknown [14]. The time delay, denoted as $\delta t$, was calculated based on Equation 1.

$$\delta t = t_{detect} - t_0 \tag{1}$$

Where $t_0$ is the time point when the experiment started and $t_{detect}$ is the time point when the global threshold was exceeded the first time after the start of experiment. The results are shown in Figure 6.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 June 2023 doi:10.20944/preprints202306.0588.v1



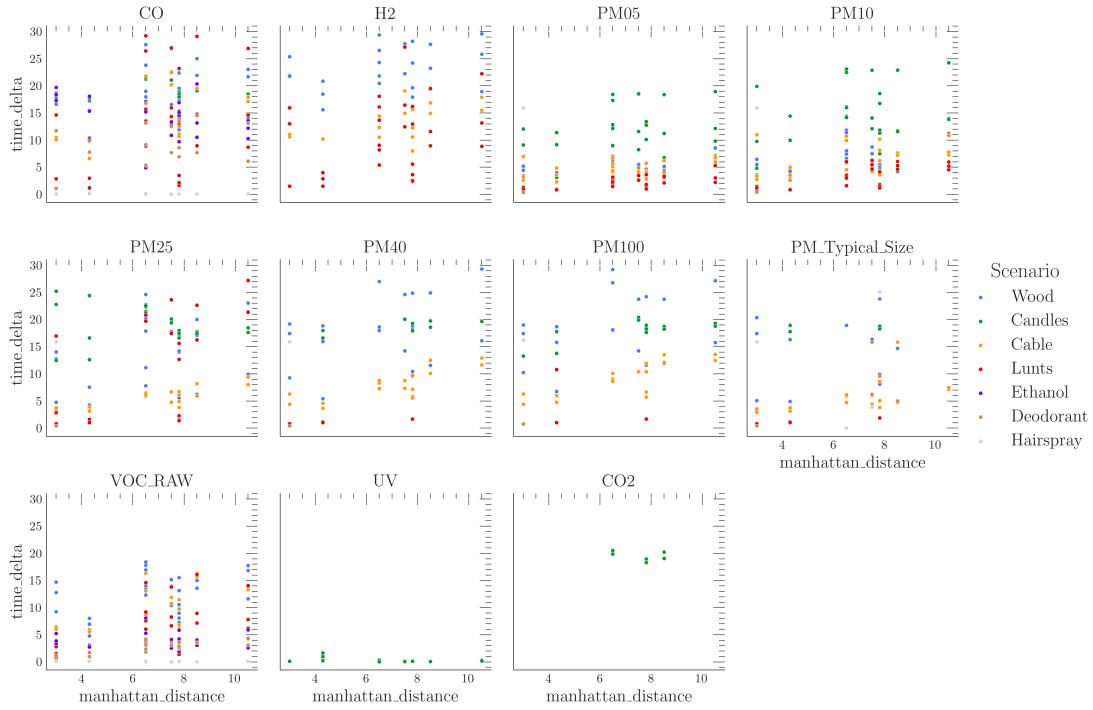


**Figure 6.** Time delta (in min) between start of experiment and threshold exceeding for each sensor measurement in dependence of the Manhattan distance to the source and the scenario.

In all experiments, the thresholds for humidity and temperature measurements were not surpassed, which is why they are not visible in Figure 7. As mentioned in Section 4.1, the maximum values of these two measurements did not exhibit any correlation with the distance from the source. This finding confirms that the carried out experiments had minimal impact on humidity and air temperature in general which contradicts the findings of Wu et al. [5] and Nakip et al. [16] who suggested using humidity and air temperature as relevant early fire indicators.

Another noteworthy observation is that the thresholds for UV and $CO_2$ measurements were only exceeded in the Candles scenario, which exclusively involved flaming combustion and a complete combustion process.

For the lunts and cable scenarios, we observed a consistent increase in time delay as the Manhattan distance increased for all measurements, except for UV.

In the wood scenario, there was a significant time delay before the threshold for $H_2$ was surpassed. In fact, it took up to half an hour from the start of the experiment for the sensor nodes closest to the emission source to exceed the threshold. This delay was considerably longer compared to the cable and lunts scenarios.

The observed delay can be attributed to the exclusive production of $H_2$ during the glowing process [27], which occurs towards the end of each wood experiment. The formation of $H_2$ is a consequence of the water gas shift reaction described in Equation 2 [28].

$$CO + H_2O \rightarrow H_2 + CO_2 \tag{2}$$

In the lunts scenario, the glowing process commenced immediately after the start of the experiment, resulting in the shortest time delta (ranging from 2 to 9 minutes, depending on the Manhattan distance) for the $H_2$ threshold to be exceeded.

In the cable scenario, the time delta for $H_2$ falls between that of the lunts and wood scenarios, ranging from 10 to 18 minutes. This can be attributed to the specific experimental procedure. In the lunts scenario, the glowing process begins right at the start of the experiment, leading to the shortest time delta. In the wood scenario, the wood initially undergoes carbonization and only starts to glow

 doi:10.20944/preprints202306.0588.v1



towards the end of the experiment, resulting in the longest time delta. In the cable scenario, the wire of the cable heats up due to electrical overload, leading to evaporation and carbonization of the insulation. This triggers the glowing process and the release of $H_2$. Although the glowing process starts earlier than in the wood scenario, it does not occur immediately at the beginning of the experiment as in the lunts scenario.

Figure 6 demonstrates that the time delta for exceeding the CO threshold in the wood scenario is similar to that of the $H_2$ measurement, albeit slightly shorter. This suggests that the presence of a significant amount of CO facilitates the release of $H_2$ during the glowing process, as described by the reaction in Equation 2.

When examining PM (Particulate Matter), it is evident that smaller particles (PM05 to PM100) in the wood scenario exhibited shorter time deltas compared to larger particles (PM25 to PM100). Conversely, in the lunts scenario, PM40 and PM100 did not surpass the threshold beyond a Manhattan distance of 8 meters. However, in the wood scenario, these particles did exceed the threshold, albeit with longer time deltas ranging from 16 to 29 minutes. This phenomenon can be attributed to the use of an external heat source in the wood experiments, as discussed in Section 3.2. The external heat source likely generated stronger thermal buoyancy, enabling the larger and heavier particles to be transported further away from the emission source in the wood experiments compared to the lunts experiments.

In the lunts scenario, VOC exhibited shorter time deltas compared to the larger particles (>PM10), but slightly longer time deltas compared to the smaller particles (PM05-PM10). However, the time deltas for VOC were still lower compared to the measurements of CO and $H_2$, and the time deltas between measurements increased as the Manhattan distance increased.

The Candles scenario stands out due to its complete combustion nature. Only during the transition to wax firing does enough $CO_2$ get produced to surpass the threshold, which explains the time delta of approximately 20 minutes. Carbon monoxide (CO) appears to be unaffected by the Manhattan distance in terms of reaching maximum values and exceeding the threshold. Hydrogen gas ($H_2$) does not occur in this particular fire scenario since there is no underlying glowing process.

In relation to PM, PM05 reaches the threshold at the fastest rate. For all sizes of PM, the time delta increases with greater Manhattan distance. Unlike the lunts scenario, the larger particles (PM40 to PM100) surpass the threshold for all Manhattan distances.

UV measurements exhibit the quickest response time among all sensors and remain unaffected by the Manhattan distance. VOC does not surpass the threshold at all.

In the cable scenario, both $H_2$ and CO took significantly longer to reach the thresholds compared to all sizes of PM, regardless of the Manhattan distance. Within the PM measurements, smaller particles (PM05) reached the threshold slightly earlier than larger particles (PM40 to PM100) at the same Manhattan distance to the source.

## 5. Discussion

Regarding to the non-gas early fire indicators, we found that UV and PM05 are least affected by the distance to between the sensor and the source within an EN54 standard room size. They are therefore the most robust features against propagation effects influencing the detection behaviour.

However, it is important to note that UV is only present in situations involving flaming combustion, which is not the case for the majority of incipient fire scenarios. [29] Therefore, we suggest employing UV not directly for ML model construction, but rather as an additional rule that can effectively verify fire detection in instances of flaming combustion.

The concentrations of larger particles (>PM10) show a significant higher dependency on the distance between sensor and source which is reflected in significant longer detection times. This is why we propose using only smaller particles (PM05 and PM10) as early fire indicators looking at the particulate matters.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 June 2023 doi:10.20944/preprints202306.0588.v1



When it comes to VOC´s, we found that they show a notable dependence on the distance to the source. They occur in all fire scenarios accept the candles fire. Since we deteced comparable concentrations in all nuisance scenarios, we do not consider VOCs to be a stand-out early fire indicator. However, as for $H_2$, in combination with PM05 and CO, VOCs can serve as a suitable early fire indicator.

Analyzing the gas-based early fire indicators, we can deduce that $CO_2$ is only significant in situations of nearly complete combustion. In our study, this was observed exclusively in the candles scenario. Therefore, we can conclude that $CO_2$ is not a relevant early fire indicator, despite its near independence from the distance between the source and the sensor in the case of a wax fire.

CO was emitted in significant quantities across all fire scenarios and exhibits only a minimal dependency on the Manhattan distance. Therefore, we highly recommend considering CO as an essential early fire indicator.

Although $H_2$ was only released in fire scenarios involving glowing processes, such as the wood and lunts experiments, we propose it as a significant early fire indicator. This is because glowing processes are a common occurrence in the majority of incipient fires [30]. The timing of $H_2$ release varies depending on the specific fire scenario, as discussed in Section 4.2. As a result, $H_2$ can be relevant as an early fire indicator either at the very beginning of an incipient fire (as seen in the lunts scenario) or after a certain period of fire development (as seen in the wood scenario).

When used in conjunction with CO, $H_2$ as an early fire indicator can enable delimitation to nuisance scenarios, as $H_2$ was not detected in all nuisance tests. However, cross-sensitivities of the CO sensor to VOCs led to a significant increase in CO. Given that the detection times for CO and $H_2$ were similar in their respective fire scenarios, the simultaneous presence of these two measured variables can be considered a reliable feature for early fire detection.

In general, it can be concluded that gas-based sensor measurements and smaller particles (PM05) are less affected by the distance between the sensor node and the source compared to larger particles (>PM05) and VOC. This characteristic makes them less reliant on the sensor's physical location and more robust as early fire indicators.

However, it should be noted that PM05-PM25 and VOC exhibit the shortest detection times, particularly in smoldering fire scenarios involving wood, cables, and lunts. Although these early fire indicators are more sensitive to the sensor node's position, they become particularly significant during the initial phase of the investigated incipient fire scenarios in this study.

## 6. Conclusion and Outlook

This paper presents the results of a series of laboratory experiments conducted with the aim of generating a comprehensive fire data set that encompasses various incipient fire scenarios in a standard EN54 indoor test room. The primary objective of these experiments was to create a new indoor fire data set and perform a study to assess the viability of different sensor measurements in terms of early fire indicators. This serves as an initial feature selection process for the development of machine learning-based models aimed at detecting the early stages of incipient fires.

We observed a correlation between the Manhattan distance from the source to the sensor node and the response behavior of specific sensor measurements in various incipient fire scenarios. In summary, we found that there is an increase in the time delay of reaching a global threshold resp. a decrease of the measured maximum with increasing Manhattan distance. However, this correlation is specific to the scenario and the sensor measurement itself.

In summary, our study identified five significant early fire indicators using a multi-sensor approach: $H_2$, CO, PM05, PM10, and VOC. The relevance of these indicators varies depending on the type of incipient fire scenario, as they become pertinent at different stages of the incipient fire. The correlations presented in this paper can be incorporated as domain knowledge into corresponding models. Furthermore, the knowledge of intensity dependence can enhance the interpretability of model predictions, allowing for a better understanding of how model performance changes at different

 doi:10.20944/preprints202306.0588.v1



positions. Moreover, the positional variations captured in this data set can be utilized to develop more robust models and mitigate the issue of overfitting.

The labeled multivariate time series data offers the opportunity to derive additional features as early fire indicators based on model assumptions. For instance, trend features can be extracted from specific subsequences of the time series, as proposed in Wu et al. [5], which forms part of ongoing work. Furthermore, since the data includes timestamps, stream-based approaches can also be modeled. The provided data set allows for training and validation of models across different sensor node positions in the room. Additionally, as each fire experiment was repeated three times, the data set provides an ample number of data points, allowing for the development and evaluation of more complex models compared to previously available data sets.

**Author Contributions:** Conceptualization, P.V. and J.K.; methodology, J.K.; software, P.V.; validation, P.V.; formal analysis, P.V.; investigation, P.V. and J.K.; resources, P.V., J.K. and S.M.; data curation, P.V.; writing—original draft preparation, P.V.; writing—review and editing, J.K. and S.M.; visualization, P.V.; supervision, U.K.; project administration, U.K.; funding acquisition, U.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the German Federal Ministry of Education and Research as part of the "Research for Civil Security" program (funding codes: 13N15415 to 13N15420 and 13N15565).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The dataset is available in CSV format and can be accessed publicly from the Mendeley Data platform at: Reserved DOI: 10.17632/npk2zcm85h.1

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Data set Description

**Table A1.** Description of the columns in the data set.

| Column | Description | Format | Unit |
|---|---|---|---|
| Date | Index column | DatetimeIndex | 'YYYY-MM-DD hh:mm:ss' |
| Sensor_ID | Unique sensor ID | string | [-] |
| CO2_Room | Concentration of carbon dioxide | float64 | ppm |
| CO_Room | Concentration of carbo monoxide | float64 | ppm |
| H2_Room | Concentration of hydrogen | float64 | ppm |
| Humidity_Room | | float64 | % |
| PM05_Room | Particles < 0.5 µm | float64 | $cm^{-3}$ |
| PM10_Room | Particles < 1.5 µm | float64 | $cm^{-3}$ |
| PM25_Room | Particles < 2.5 µm | float64 | $cm^{-3}$ |
| PM40_Room | Particles < 4.0 µm | float64 | $cm^{-3}$ |
| PM100_Room | Particles < 10.0 µm | float64 | $cm^{-3}$ |
| PM_Room_Typical_Size | Weighted mean of diameter | float64 | µm |
| Temperature_Room | Air temperature | float64 | °C |
| UV_Room | UV photon counts | float64 | # |
| VOC_Room_RAW | Volatile organic compounds (raw electrical data from sensor) | float64 | A.U. |
| scenario_label | experiment specific label | string | [-] |
| anomaly_label | distinguishes between 'Anomaly' and 'Normal' | string | [-] |
| ternary_label | distinguishes between 'Nuisance', 'Fire' and 'Background' | string | [-] |

## References


1. Tsujita, W.; Yoshino, A.; Ishida, H.; Moriizumi, T. Gas sensor network for air-pollution monitoring. *Sensors and Actuators B: Chemical* **2005**, *110*, 304–311. doi:10.1016/j.snb.2005.02.008.
2. Fireteanu, V.V.; Tudor-Serban, A.; Sacala, I.S.; Moisescu, M.A. Avalanche Prediction Based on Snow Level Monitoring Using Wireless Sensor Networks. *Applied Mechanics and Materials* **2014**, *656*, 369–377. Conference Name: Monitoring, Controlling and Architecture of Cyber Physical Systems ISBN: 9783038352747 Publisher: Trans Tech Publications Ltd, doi:10.4028/www.scientific.net/AMM.656.369.
3. Mesquita, E.; Arêde, A.; Pinto, N.; Antunes, P.; Varum, H. Long-term monitoring of a damaged historic structure using a wireless sensor network. *Engineering Structures* **2018**, *161*, 108–117. doi:10.1016/j.engstruct.2018.02.013.
4. Dinh, T.L.; Hu, W.; Sikka, P.; Corke, P.; Overs, L.; Brosnan, S. Design and Deployment of a Remote Robust Sensor Network: Experiences from an Outdoor Water Quality Monitoring Network. 32nd IEEE Conference on Local Computer Networks (LCN 2007), 2007, pp. 799–806. ISSN: 0742-1303, doi:10.1109/LCN.2007.39.
5. Wu, L.; Chen, L.; Hao, X. Multi-Sensor Data Fusion Algorithm for Indoor Fire Early Warning Based on BP Neural Network. *Information* **2021**, *12*, 59. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/info12020059.
6. Rachman, F.Z.; Hendrantoro, G.; Wirawan. A Fire Detection System Using Multi-Sensor Networks Based on Fuzzy Logic in Indoor Scenarios. 2020 8th International Conference on Information and Communication Technology (ICoICT), 2020, pp. 1–6. doi:10.1109/ICoICT49345.2020.9166416.
7. Liang, Y.h.; Tian, W.m. Multi-sensor Fusion Approach for Fire Alarm Using BP Neural Network. 2016 International Conference on Intelligent Networking and Collaborative Systems (INCoS), 2016, pp. 99–102. doi:10.1109/INCoS.2016.38.
8. Jana, S.; Shome, S.K. Hybrid Ensemble Based Machine Learning for Smart Building Fire Detection Using Multi Modal Sensor Data. *Fire Technology* **2022**. doi:10.1007/s10694-022-01347-7.
9. Yu, M.; Yuan, H.; Li, K.; Wang, J. Research on multi-detector real-time fire alarm technology based on signal similarity. *Fire Safety Journal* **2022**, p. 103724. doi:10.1016/j.firesaf.2022.103724.
10. Milke, J.A.; Mcavoy, T.J. Analysis Of Fire And Non-fire Signatures For Discriminating Fire Detection. *Fire Safety Science* **1997**, *5*, 819–828.
11. Solórzano, A.; Eichmann, J.; Fernández, L.; Ziems, B.; Jiménez-Soto, J.M.; Marco, S.; Fonollosa, J. Early fire detection based on gas sensor arrays: Multivariate calibration and validation. *Sensors and Actuators B: Chemical* **2022**, *352*, 130961. doi:10.1016/j.snb.2021.130961.
12. Andrew, A.M.; Shakaff, A.; Zakaria, A.; Gunasagaran, R.; Kanagaraj, E.; Saad, S.M. Early Stage Fire Source Classification in Building using Artificial Intelligence. 2018 IEEE Conference on Systems, Process and Control (ICSPC), 2018, pp. 165–169. doi:10.1109/SPC.2018.8704155.
13. Bukowski, R.W.; Peacock, R.D.; Averill, J.D.; Cleary, T.G.; Bryner, N.P.; Reneke, P.A. Performance of Home Smoke Alarms, Analysis of the Response of Several Available Technologies in Residential Fire Settings. *NIST* **2003**. Last Modified: 2017-02-19T20:02-05:00 Publisher: Richard W. Bukowski, Richard D. Peacock, Jason D. Averill, Thomas G. Cleary, Nelson P. Bryner, Paul A. Reneke.
14. Nazir, A.; Mosleh, H.; Takruri, M.; Jallad, A.H.; Alhebsi, H. Early Fire Detection: A New Indoor Laboratory Dataset and Data Distribution Analysis. *Fire* **2022**, *5*, 11. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/fire5010011.
15. Gottuk, D.T.; Peatross, M.J.; Roby, R.J.; Beyler, C.L. Advanced fire detection using multi-signature alarm algorithms. *Fire Safety Journal* **2002**, *37*, 381–394. doi:10.1016/S0379-7112(01)00057-1.
16. Nakıp, M.; Güzeliş, C. Multi-Sensor Fire Detector based on Trend Predictive Neural Network. 2019 11th International Conference on Electrical and Electronics Engineering (ELECO), 2019, pp. 600–604. doi:10.23919/ELECO47770.2019.8990400.
17. Chen, S.J.; Hovde, D.C.; Peterson, K.A.; Marshall, A.W. Fire detection using smoke and gas sensors. *Fire Safety Journal* **2007**, *42*, 507–515. doi:10.1016/j.firesaf.2007.01.006.
18. Gutmacher, D.; Hoefer, U.; Wöllenstein, J. Gas sensor technologies for fire detection. *Sensors and Actuators B: Chemical* **2012**, *175*, 40–45. doi:10.1016/j.snb.2011.11.053.

19. Guo, S.; Yang, R.; Zhang, H.; Zhang, X. New Inverse Model for Detecting Fire-Source Location and Intensity. *Journal of Thermophysics and Heat Transfer* **2010**, *24*, 745–755. Publisher: American Institute of Aeronautics and Astronautics _eprint: https://doi.org/10.2514/1.46513, doi:10.2514/1.46513.
20. Kou, L.; Wang, X.; Guo, X.; Zhu, J.; Zhang, H. Deep learning based inverse model for building fire source location and intensity estimation. *Fire Safety Journal* **2021**, *121*, 103310. doi:10.1016/j.firesaf.2021.103310.
21. Fonollosa, J.; Solórzano, A.; Marco, S. Chemical Sensor Systems and Associated Algorithms for Fire Detection: A Review. *Sensors* **2018**, *18*, 553. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/s18020553.
22. Alessandri, A.; Bagnerini, P.; Gaggero, M.; Mantelli, L. Parameter estimation of fire propagation models using level set methods. *Applied Mathematical Modelling* **2021**, *92*, 731–747. doi:10.1016/j.apm.2020.11.030.
23. Mostafa, E.; Nannen, C.; Henseler, J.; Diekmann, B.; Gates, R.; Buescher, W. Physical properties of particulate matter from animal houses—empirical studies to improve emission modelling. *Environmental Science and Pollution Research* **2016**, *23*, 12253–12263. doi:10.1007/s11356-016-6424-8.
24. Bin, H.; Yang, Y.; Cai, L.; Zhulin, Y.; Roszak, S.; Linjun, Y. Experimental study on particles agglomeration by chemical and turbulent agglomeration before electrostatic precipitators. *Powder Technology* **2018**, *335*, 186–194. doi:10.1016/j.powtec.2018.04.016.
25. Evtyugina, M.; Alves, C.; Calvo, A.; Nunes, T.; Tarelho, L.; Duarte, M.; Prozil, S.O.; Evtuguin, D.V.; Pio, C. VOC emissions from residential combustion of Southern and mid-European woods. *Atmospheric Environment* **2014**, *83*, 90–98. doi:10.1016/j.atmosenv.2013.10.050.
26. Kim Oanh, N.T.; Ly, B.T.; Tipayarom, D.; Manandhar, B.R.; Prapat, P.; Simpson, C.D.; Sally Liu, L.J. Characterization of particulate matter emission from open burning of rice straw. *Atmospheric Environment* **2011**, *45*, 493–502. doi:10.1016/j.atmosenv.2010.09.023.
27. Cofer III, W.R.; Winstead, E.L.; Stocks, B.J.; Goldammer, J.G.; Cahoon, D.R. Crown fire emissions of CO2, CO, H2, CH4, and TNMHC from a dense Jack pine boreal forest fire. *Geophysical Research Letters* **1998**, *25*, 3919–3922. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/1998GL900042, doi:10.1029/1998GL900042.
28. Kohl, D.; Kelleter, J.; Petig, H. Detection of Fires by Gas Sensors. *Sensors Update* **2001**, *9*, 161–223. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/1616-8984%28200105%299%3A1%3C161%3A%3AAID-SEUP161%3E3.0.CO%3B2-A, doi:10.1002/1616-8984(200105)9:1<161::AID-SEUP161>3.0.CO;2-A.
29. Kropotova, S.S.; Kuznetsov, G.V.; Strizhak, P.A. Identifying products of pyrolysis and combustion of materials at incipient stages of fires. *Fire Safety Journal* **2022**, *132*, 103643. doi:10.1016/j.firesaf.2022.103643.
30. Vasiliev, A.A.; Grigoriev, G.Y.; Lagutin, A.S.; Nabiev, S.S.; Pisliakov, A.V.; Samotaev, N.N.; Sokolov, A.V. Contemporary technologies of early detection of fire in space vehicles. *Acta Astronautica* **2017**, *135*, 76–82. doi:10.1016/j.actaastro.2016.11.005.