SEA-HAZEMON: Active Haze Monitoring and Forest Fire Detection Platform

Adisorn Lertsinsrubtavee
intERLab
Asian Institute of Technology
Pathumthani, Thailand
adisorn@ait.ac.th

Kalana G.S. Jayarathna
intERLab
Asian Institute of Technology
Pathumthani, Thailand
kalana@ait.ac.th

Preechai Mekbungwan
intERLab
Asian Institute of Technology
Pathumthani, Thailand
preechaim@ait.ac.th

Thongchai Kanabkaew
Faculty of Public Health
Thammasat University
Pathumthani, Thailand
thongchai.k@fph.tu.ac.th

Sunee Raksakietisak
Thai-Australian Technological
Services Center
Bangkok, Thailand
sunee@g.swu.ac.th

ABSTRACT

Every year, South East Asian (SEA) region has suffered from haze pollution significantly caused by forest fires and agricultural-related burning. Mitigating this problem requires a robust system that can monitor haze and air pollution in real time across the region. The SEA-HAZEMON is an IoT platform that consists of low-cost air quality sensors and several cloud services. The platform also consists of a forest fire detection model based on particulate matter (PM2.5) and Carbon Monoxide (CO) concentration. In our trial, the early warning messages were timely sent to navigate the local forest fire authorities via short messaging applications (i.e., Telegram and Line). The notification messages were analyzed together with real forest fire incidents that occurred in the northern part of Thailand during the fire burning season. The results showed that there were 367 fire events detected in April 2022 that achieved an accuracy of 84%. Our study also discovered the effect of humidity which reduced the accuracy of our model by approximately 13%.

CCS CONCEPTS

• **Computer systems organization → Embedded and cyber-physical systems**.

KEYWORDS

Internet of Things, low-cost micro sensor, forest fire detection, wildfire, PM2.5, haze monitoring

ACM Reference Format:
Adisorn Lertsinsrubtavee, Kalana G.S. Jayarathna, Preechai Mekbungwan, Thongchai Kanabkaew, and Sunee Raksakietisak. 2022. SEA-HAZEMON: Active Haze Monitoring and Forest Fire Detection Platform. In *The 17th Asian Internet Engineering Conference (AINTEC'22), December 19–21, 2022, Hiroshima, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3570748.3570761

1 INTRODUCTION

The massive forest fire outbreaks during the past decade have exacerbated global climate change and health impacts in various parts of the world. Forest fire and biomass burning are one of the major sources to generate small particulate matter (PM2.5) emissions into the atmosphere [14]. Once inhaled, these small particles can penetrate deeply into the heart and lungs causing several respiratory diseases. To mitigate this issue, providing an early warning to prevent large forest fire damage is very crucial. Traditionally, the local forest fire authorities and volunteers have to regularly monitor the fire situation by patrolling around the forest and national reserved areas. However, this method is limited to large-area monitoring which is hard to provide timely information. For wider observation, a satellite-based system, namely Fire Information for Resource Management System (FIRMS) has been used to monitor global fire events [13] for more than a decade. The near real-time hotspot locations are periodically collected through the Moderate Resolution Imaging Spectroradiometer (MODIS) equipped on Terra and Aqua satellites, which can report active fire maps worldwide. However, the satellite data is still lacking precision and fine resolution on ground measurement [10]. Besides, with a fixed schedule of the satellite orbit, the fire hotspot measurement is captured only once or twice a day. On the other hand, commercial forest fire detection instruments such as thermal cameras and precision smoke detectors have been widely proposed, but their cost is very expensive (i.e., 10,000 - 30,000 USD per unit) [3]. In addition, setting up those tools typically is limited to a stable installation site which requires reliable electric power and network connection.

Alternatively, the emergence of the Internet of Things (IoT) enables the use of new low-cost and tiny sensors for real-time environmental monitoring applications. Electronic sensors like PM2.5, CO, CO_2 and O_3 have been widely integrated into IoT devices to indicate the start of a potential fire incident [9, 17]. Among those sensors, PM2.5 with light scattering detection technique is the most acceptable one due to its accuracy and high correlation with the reference instruments [1]. Nevertheless, the use of IoT devices is still considered only in the laboratory setting and urban/semi-urban

AINTEC'22, December 19–21, 2022, Hiroshima, Japan

ACM ISBN 978-1-4503-9981-4/22/12…$15.00
https://doi.org/10.1145/3570748.3570761

areas, the practical deployment in a challenging environment like a forest area has not been much explored.

To better observe and prevent the impact of forest fire occurrences, this paper proposes a robust air pollution monitoring system that can collect data in a remote area (i.e., forest area, agricultural field). A low-cost IoT device, called Canarin is developed with the following objectives; providing both online and offline measurements of PM2.5 and other related parameters (CO, CO_2, RH, etc.), powered by solar and rechargeable battery and capable of long-term deployment. The cloud platform, called SEA-HAZEMON is developed to support large-scale data collection, real-time monitoring, and big data analytics for detecting forest events. The platform is also integrated with short messaging services (e.g., Line and Telegram) to provide timely forest fire warnings. Consequently, a network of Canarin nodes was deployed in Doi Chang Pa Pae, the highest mountain in Lamphun province, Thailand, to monitor the forest fire. Specifically, our key contributions are summarized as follows:

- We share our firsthand experience in successfully designing and building a real-world field air quality sensors network in a deep forest area with low-cost IoT devices. These devices have proven to be a self-sustain system for long-term data observatories.
- Analysis of the data collected during the forest fire events provides timely alert messages to local authorities in preventing substantial damage.
- We observe that relative humidity (RH) factor plays a key role to distort the efficiency of sensor readings. Through extensive data analysis, we shed light on how to filter some erratic data while improving the accuracy of forest fire detection model. Besides, we suggest an effective area and point out suitable location for deployment.

2 HAZE AND FOREST FIRE MONITORING PLATFORM

This section presents the SEA-HAZEMON platform that aims to monitor and detect forest fire events in real time. The platform contains two parts; 1) a low-cost sensor node to gather data collection in the field and 2) a cloud back-end system that is responsible for data storage and data visualization.

2.1 Low-Cost Haze Monitoring Sensor

A low-cost haze monitoring sensor node, called Canarin has been developed under the SEA-HAZEMON project [15]. A Canarin sensor node is built from the UDOO Neo, a single board computer that contains two CPU cores: ARM Cortex-A9 and Cortex M4 which run both Ubuntu Linux distribution and full stack Arduino environment at the same time. The Arduino part and onboard pin connectors are compatible with most of the sensors and actuators. Figure 1 illustrated the blueprint of Canarin node that integrates with multi-sensors including PMS 7003 (light scattering PM1/2.5/10 sensor), BME 280 (relative humidity, temperature, and air pressure sensor), ZE-07 (CO sensor), MH-Z16 (CO_2 sensor) and Ublox M8N (GPS module).

On the other hand, the Linux core acts as a central control unit for collecting and transmitting data to the cloud. A python script, called *Ardu2Linux* handles the sensor data readings from the Arduino core through the UART interface. Each collected data is attached with GPS location and timestamp and stored on a local SD card. The timestamp can be updated from various synchronization sources including a preconfigured NTP server, GPS information, and a built-in RTC module. The *Linux2Server* script is responsible for reading the data from local storage and creating UDP packets encoded in Type-Length-Value (TLV) format. The Canarin's data packet will be transmitted to the pre-configured address: *hazemon.in.th* via a WiFi connection. To provide reliable transmission, we follow a simple stop-and-wait protocol by waiting for an acknowledgment from our server before sending a new packet. After receiving the acknowledgment, the data packet will be deleted from the memory. The re-transmission will proceed if a node could not receive the acknowledgment before timeout. In case, a node could not connect to the server (i.e., an internet connection is not available), the data packet will be kept in the local SD card. When the connection is resumed, the pending packets in the queue will be gradually sent to the server.

Our goal is to deploy the Canarin node in the deep forest area where a stable electric source is not available. In this regard, a self-harvesting solar station was designed and built to provide an alternative power source for the Canarin node. Each station consists of a solar panel, a solar charge controller, a Lithium-ion battery, a circuit breaker, and a 4G WiFi router for internet connectivity. A waterproof box was modified to enclose the Canarin node and other equipment as shown in Figure 2. Table 1 illustrates a total power consumption which is about 330 mA/ 3.9 Wh. Based on this calculation, we select the battery size as 25 Ah with a 12 V supply. Theoretically, our devices could last for 76 hours (more than 3 days) with a fully charged battery. As for the solar panel, a 100 Wh harvesting capability is chosen where an empty battery could be fully charged within 3 hours. The specification is designed to compensate for the efficiency dropped during the deployment. Given that the deployed location is inside a tropical forest surrounded by thick trees, the efficiency of a solar panel can be reduced dramatically. However, we assume that a solar panel could have at least 3 hours to harvest the energy, then our Canarin could continuously operate without interruption.

Table 1: List of equipment and average power consumption

Equipment	Volt	mA	Wh
Canarin Sensor	12	300	3.6
4G/WiFi Router	9	20	0.18
Solar charge controller	12	10	0.12
Total		330 mA	3.9 Wh

Prior to the field deployment, each Canarin node was tested in our laboratory for reading consistency among different nodes by putting all nodes together under the same environment. All nodes were found to give similar readings according to a sensor datasheet. From this process, the sensor nodes that happened to report different readings from the rest of the group were excluded and fixed. As a consequence, the Canarin sensor nodes were also

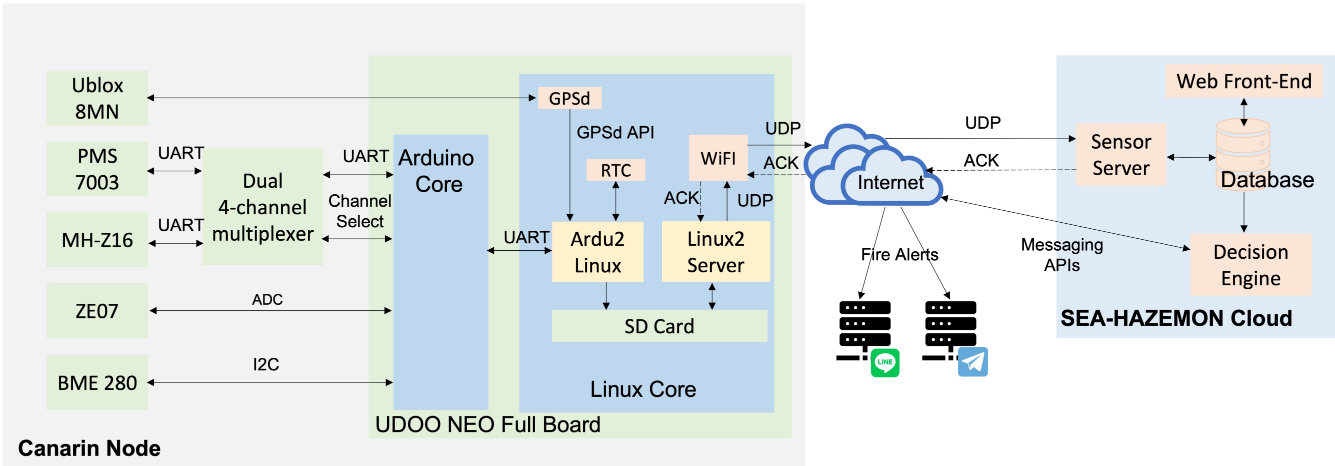
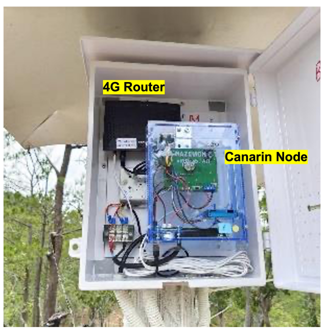
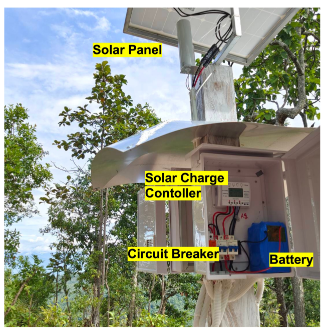


Figure 1: The architecture Canarin sensor node and SEA-HAZEMON cloud platform



(a) A Canarin sensor node and 4G/WiFi router are kept in the enclosure box.



(b) The solar charging circuit and battery are stored in another enclosure box.

Figure 2: A Canarin sensor node and solar power station deployed in Doi Chang Pa Pae Mountain.

Figure 3: The Canarin nodes were tested in an ambient environment by co-locating with a Beta Rey air quality station

tested in an ambient environment by co-locating them at the reference air quality station owned by the Pollution Control Department of Thailand (PCD). The station locates in the downtown area of Bangkok close to the Rama-IV road in the Pathumwan district. The station uses the beta ray attenuation method for measuring the PM2.5 concentration in an hourly average. Figure 4 illustrates an experimental setup in an ambient environment where the Canarin nodes were temporarily set up on the station roof (See Figure 4). As shown in Figure 4 and 5, PM2.5 reading from Canarin nodes correlated well with the reference instrument with the coefficient of determination (R^2) of 0.64.

2.2 SEA-HAZEMON Cloud

The SEA-HAZEMON offers several services running over the cloud network. Those services include a web front end for data visualization, a database, and an active notification system. The architecture of the SEA-HAZEMON cloud is illustrated in Figure 1 which contains four main components as follows:

Sensor Server collects data stream from each online Canarin sensor node. The communication between a sensor node and the sensor server is based on the traditional UDP protocol where each node transmits its data through a WiFi connection. Once this data is received, the payload is extracted and inserted into the sensor database. The acknowledgment message is used to confirm a successful transmission.

Cloud Data Storage is based on the MySQL database. Each data consists of a unique node ID, GPS location of the sensor node (latitude, longitude, and altitude), sensor type, sensor value, and the sampling timestamp in the form of key/value pairs allowing data to be inserted into the MySQL database irrespective of a node ID.

Web Front End is in charge of data visualization. It periodically retrieves data from the MySQL database. This web front end provides a user interface for internet users to visualize the status of haze monitoring in real time. In addition, it also provides an

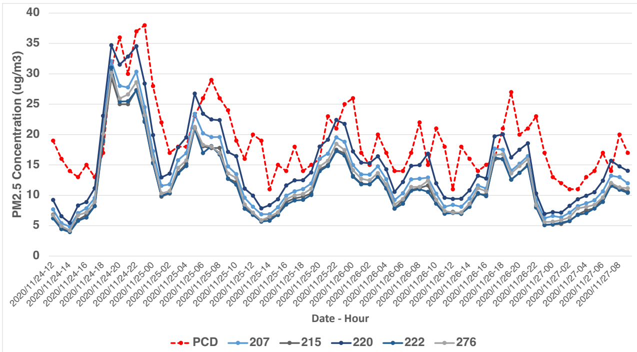


Figure 4: The time series plots of PM2.5 concentration collected from the Canarin nodes and the Beta Rey air quality station station (PCD).

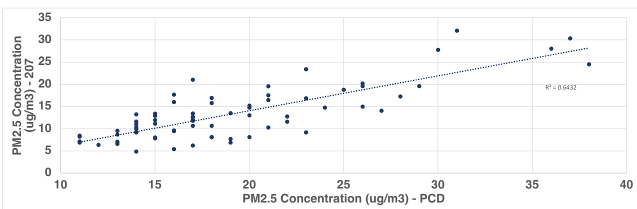


Figure 5: The scatter plot comparing the PM2.5 collected from a Canarin node number 207 and the Beta Rey air quality station station (PCD)

interface for downloading the open data in CSV format for further data analytics propose.

Decision Engine (DE) is a core component that can make the strategic decision of critical alerts. Several algorithms can be further deployed for different decision-making on haze pollution and other related applications. The DE is also connected with short messaging services for sending push notifications to our subscribed users. Currently, SEA-HAZEMON has connected to Line and Telegram through Messaging API. The official channels on Telegram and Line were created to receive updates and notification messages from SEA-HAZEMON services.

2.3 Forest Fire Detection Model

The forest fire detection model has been integrated into the Decision Engine for detecting forest fire events in real-time. The model was developed from our previous work [15] using a non-parametric supervised learning method with the forest fire observation and a sensor dataset in 2021. The algorithm was derived from a decision tree model that classifies and predicts the forest fire event based on PM2.5 and CO concentrations as presented in Algorithm 1.

At the first step, the decision engine retrieves the PM2.5 and CO samplings over the past T period. In our case, T is set as 15 minutes following the suggestion in [7]. Then, it calculates an average value of PM2.5 concentration for each sensor node while comparing it with PM2.5 threshold values. The algorithm is classified into three states based on average PM2.5 concentration. At the *High*

Algorithm 1 Forest Fire Detection

```
Input PM_i(T): PM_2.5 values of node i in period T , CO_i(T):
  CO values of node i in period T, S: Set of sensor nodes
for all i ∈ S do
  p ← Average(PM_i(T)) {Average PM values of T}
  c ← Max(CO_i(T)) {Maximum CO values of T }
  if p > 121.22 {High State }
    Send Notification ( )
  else if   71.22 < p < 121.22 {Mid State }
    if c > 0
     Send Notification ( )
  else {Low State }
    return No fire detected
end for
```

state, the average PM2.5 concentration is higher than 121.22 μ/m^3 which is classified as burning (i.e., a forest fire event is detected). Consequently, the *Moderate* state means the PM2.5 concentration is between 121.22 and 71.22 μ/m^3. The burning is detected if the *CO* concentration is higher than 0 ppm. Lastly, if the $PM_{2.5}$ is less than 71.22, the state is determined as *Low* where the burning is not detected. The function *Send Notification ( )* is called when the burning is detected. Consequently, the alert message will be sent to both users subscribed to our Telegram and Line official accounts.

3 SYSTEM DEPLOYMENT

The Doi Chang Pa Pae area was chosen for our study which is the highest mountain located at the border between Lamphun and Chiangmai provinces in northern Thailand. The local community living in this area is typically the Karen Sagor hill tribe who are doing mountain agriculture for their living. Many of the community's members have volunteered to monitor the forest fire situation in their area. To help them monitor the fire situation, the Canarin sensor nodes together with solar stations were deployed in 5 locations including 1) Phu Huai Pu (222), 2) Mon Wai(276), 3) Buak Tong Tung (207), 4) Buang Tum Boon (220) and 5) Ta Dedo (215) as shown in Figure 6. Each Canarin node was placed around the firebreak ring which is used as a barrier to slow down the progress of forest fire. All five Canarin sensor nodes have been operated since July 2021 using self-harvesting solar energy. An example of deployed Canarin sensor node at the Buang Tum Boon location (215) with a solar station is presented in Figure 2.

4 RESULTS AND DISCUSSIONS

Our main objective is to determine how the low-cost sensor nodes would perform in a forest area while focusing on the efficiency of our proposed forest fire detection model. To identify the actual forest fire events, we obtained the ground truth data from the fire report collected by Ban Hong and Jom Thong forest fire authorities together with the satellite-based hotspots from FIRMS [13]. The data were collected from 1 to 28 April 2022 which was considered as the peak of the forest fire season in Doi Chang Pa Pae. Consequently, the boundary of our study area was defined by drawing a 5 km line from each direction. For instance, the northbound was drawn from node 207, the westbound was drawn from node 276, while the

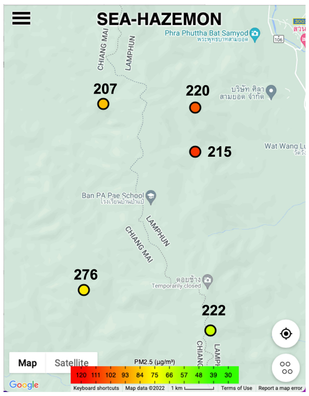


Figure 6: The location of Canarin nodes deployed in Doi Chang Pa Pae mountain

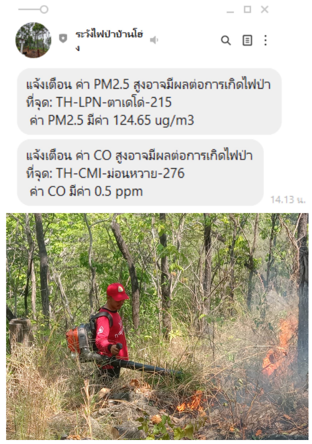


Figure 7: The local forest fire officers received the alert messages from SEA-HAZEMON and operated the fire mission on 5 April 2022.

southbound and eastbound were drawn from node 222 (see Figure 6). From the ground truth data, there were 10 events detected in our study area, their GPS coordinates (flame icon) are illustrated in Figure 8 together with the location of Canarin nodes (yellow pin).

4.1 Data Analysis on PM2.5 and CO Concentrations

This section aims to analyze the data collected from all 5 Canarin nodes to investigate the behavior of PM2.5 concentration and other related parameters. Each node was configured to record data for approximately 2 minutes per cycle while periodically uploading the data to the SEA-HAZEMON cloud. Based on our previous study [7], the 15 minutes interval was found to be suitable for haze monitoring thus the raw data were grouped and averaged for every 15 minutes interval.

The time series plots of the daily average PM2.5 and maximum CO concentration during 1 - 28 April 2022 are illustrated in Figure 9. The PM2.5 concentration was varied over a month where the highest PM2.5 concentration (153.9 $\mu g/m^3$) was occurred on 8 April 2022 and the lowest PM2.5 (less than 5 $\mu g/m^3$) was reported on 3 April 2022. On the other hand, only few nodes could capture CO gas. The highest CO concentration (8.7 ppm) was reported by node 276 on 22 April 2022, but the average PM2.5 concentration was only around 20 $\mu g/m^3$.

4.2 Analysing the Forest Fire Notifications

To predict the forest fire event, the sensor data including PM2.5 and CO parameters were collected through the SEA-HAZEMON cloud every 2 minutes while calculating the averaged value within 15 minutes intervals. If a forest fire event is detected, an alert message will be sent to the subscribed Telegram and Line users. An interesting event was reported on 5^{th} April 2022 where the forest

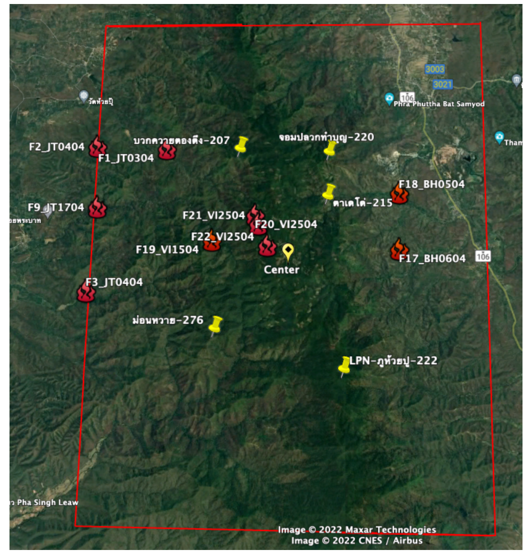


Figure 8: The locations of hotspots (flame icon) captured by the local forest fire authorities and FIRMS during 1-28 April 2022. The yellow pin icons represent the location of the Canarin sensor nodes.

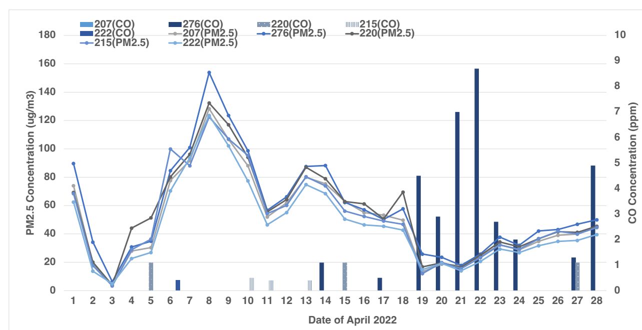


Figure 9: Time series plots of daily average PM2.5 and maximum CO concentrations during 1 - 28 April 2022

fire event was detected at the Buang Tum Boon area (node 220) and alert messages were sent to the local forest fire authorities and volunteers in Doi Chang Pa Pae. Consequently, the officers from Ban Hong forest fire station went out on a fire mission to limit the further spread of fires. The fire source was discovered 3km away from node 220. Figure 7 shows the alert messages sent out from the SEA-HAZEMON platform which were taken from the fire mission on 5 April 2022. The messages included the current level of PM2.5, CO, and node location directing to the nearest burning source.

During the study period, there were 871 events notified from all 5 sensor nodes. Prior to the data analysis, the fire events that occurred within the same 15 minutes time slice were grouped into one event. As a result, the total number of notified events was reduced to 367 events. Table 2 presents the confusion matrix comparing the predicted results from the forest fire detection model and actual events from the ground truth data. The model achieved 32 *True*

Table 2: A confusion matrix visualizes and summarizes the performance of the forest fire detection algorithm.

Actual Event	Predicted Event	
	Detected Fire	***Not Detected***
Detected Fire	True Positive (32)	False Negative (113)
Not Detected	False Positive (335)	True Negative (2303)

Positive (TP) and 2303 *True Negative (TN)* cases where the predicted results are matched with the actual events. On the other hand, our model achieved 335 *False Positive (FP)* and 113 *False Negative (FN)*. The efficiency of the forest fire detection model is evaluated in terms of recall, precision, and accuracy which can be calculated by the following equations:

$$Recall = \frac{TP}{TP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

The accuracy of the forest fire detection model is about 83.9% while achieving precision and recall as 8.71% and 22.07% respectively. The harmonic mean of both precision and recall is computed as F1-Score as the following equation.

$$F1 - Score = 2x \frac{PrecisionxRecall}{Precision + Recall} \tag{4}$$

As a result, the model achieved F1-Score of 0.125. Even though the mode is satisfied with a high accuracy rate, the precision and recall are slightly low. This is because there were many false positive and false negative cases which will be further investigated in the following section.

4.3 Effects of Humidity

In this section, we examine the humidity factor that affects the accuracy of the model. The majority of false positive cases came from the efficiency of the PM2.5 (particulate matter) sensor. The low-cost PM2.5 sensor equipped in the Canarin node is based on light scattering technology. The sensor is operated by a single light source (visible or near-infrared) that illuminates an air channel where a particle can deviate a part of light accordingly to the principle of light scattering [6]. However, this technique is sensitive to a high-humidity environment due to steam or vapor could significantly deviate the light.

To investigate this issue, the true positive (TP) and false positive (FP) cases are grouped by each day while calculating the daily average of relative humidity. The average relative humidity (RH) of true positive and false positive days are compared and illustrated in Figure 10. On false positive days, the average RH is higher than 55% while the true positive days obtain lower RH than 50%. Except for 14 April 2022, the average relative humidity is rather high (68%) even though it is considered a true positive day. Similar to the work

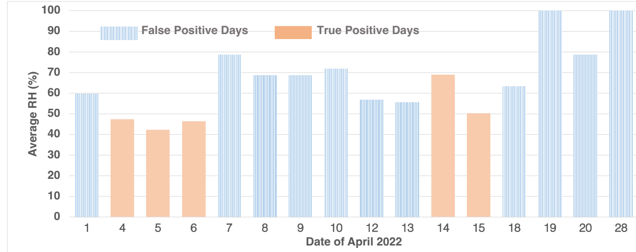


Figure 10: The bar chart compares a daily average of relative humidity on true positive and false positive days.

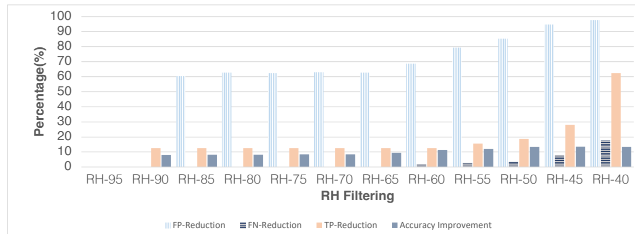


Figure 11: The bar chart illustrates the accuracy improvement of the forest fire detection model after removing the humidity factor.

in [11], the performance of low-cost PM2.5 sensors was significantly dropped when the relative humidity is higher than 80%.

With the effect of humidity, the accuracy of the forest detection model could be improved by mitigating the effects in a high-humidity environment. In this context, we apply a relative humidity threshold (RH_T) and remove some erratic samples before running the model. Figure 11 shows the improvements of our model after applying RH_T which is varied from 95% to 40%. The accuracy of our model is improved from 7.88% to 13.62%. This is due to the significant reduction of false positive (FP) cases. As shown in the plot, the number of FP cases is reduced from 60.6% to 97% when the RH_T is applied from 85% to 40%. Although the accuracy is slightly improved using RH_T filtering, the efficiency of the model is impacted by the reduction of true positive cases. As a result, the model provides incorrect results on fire detection. To balance this trade-off, an optimal threshold for RH_T must be carefully chosen by considering the reduction of true positive cases. A number of true positive (TP) cases is suddenly dropped to 12.5% when the 95% of RH_T is applied. The reduction remains 12.5% until the RH_T is increased to 60% which is considered as the maximum boundary of the relative humidity threshold. This value is also matched with analysis in Figure 10 where the average relative humidity of false positive days is higher than 55%.

4.4 Distance and Altitude

Even though the false positive cases are improved substantially when the RH_T is applied, the accuracy is still not satisfied as the number of false negative is not improved. We suspect that the location of the forest incidents from the ground truth data could be far from our sensor nodes. To verify this issue, we examine the locations of fire incidents while comparing them with our detection result either the incident was detected (TP) or missed (FN). The distance ($Dist$) and altitude difference ($Diff$) from each incident to the nearest sensor node are also measured for the analysis as presented in Table 3. Notice that the negative value of $Diff$ means the incident is located at a lower altitude than the sensor node.

As shown in the table, all true positive (TP) cases were detected within a range of 5km. Intuitively, the accuracy of our model is also related to altitude difference as most of the false negative cases were located at a higher altitude than our sensor nodes where the values of $Diff$ are negative. As a matter of fact, fire smoke is vertically blown up to a higher altitude as the air pressure is low during the day.

Table 3: Analyzing the reported locations of forest fire incidents comparing with the location of nearest sensor node

Hotspot	Alt (m)	$Diff$ (m)	$Dist$ (km)	Result
F1	559	-0.46	2.52	FN
F2	558	-467	4.99	TP
F3	517	-586	4.51	TP
F17	506	-556	3.23	TP
F18	414	-648	2.63	TP
F19	920	-183	3.05	TP
F9	624	-401	5.31	FN
F20	1086	24	2.52	FN
F21	1108	46	2.87	FN
F22	1082	20	2.66	FN

4.5 Discussions

The use of low-cost sensor nodes in the Doi Chang Pa Pae area has proven that our system can provide reliable and timely warning messages to the local forest fire authorities and volunteers. The key success factor is attributed to its self-sustainable system. Technically, the Canarin sensor node is robust that provides continuously data collection in its one-year operating time. A few records of system failures are due to running out of battery. Most glitches are caused by the lack of sunlight during the rainy season. The unstable internet connection was also another key factor that blocked the data transmission. Thanks to the robustness of our system that provides the local storage on-board SD card that can log the data for up to 3 months without an internet connection. Besides, our node also contains several scripts running in the background to protect against system failures such as checking the WiFi connection, SD card, and system load. If an undesired event is detected, the system will be rebooted automatically. These functions are very helpful as our nodes can be self-recovery and keep the system running constantly.

The evaluation in section 4.2 reveals that the forest fire detection model achieved higher accuracy with larger than 80%. However, the precision and recall values are not yet satisfied with a bit lower rate. This can be explained through a high number of false positive and false negative cases. The relative humidity factor significantly reduces the number of false positive from 335 to 1 case, if the model opted to use 40% as a cut-off threshold (RH_T). However, the false positive cases can be reconsidered as real forest fire events. As a matter of fact, all 5 sensor nodes were deployed around the firebreak ring located at the cliff or close to the top of the mountain. Practically, this area is difficult to access and could take more than 2 hours to walk from the mountain base. Determining the fire records from forest fire officers, most of the events were found at the lower altitude (414 to 624 meters), where there were around at the base of the mountain. Intuitively, the events detected by our sensor nodes may not be discovered by the forest fire officers. Regarding the analysis in subsection 4.4, a 5 km distance was observed as the effective area for a sensor node to detect the forest fire event. Therefore, deploying more nodes to cover larger areas will substantially reduce the number of false negative cases, and thus the accuracy of forest fire detection is improved.

5 RELATED WORKS

The solution for near real-time fire detection and monitoring systems has been explored for several years. The Moderate Resolution Imaging Spectroradiometer (MODIS) instruments equipped on Terra and Aqua satellites have actively provided a global fire map for more than a decade [12]. A recent higher resolution imagery instrument called VIIRS equipped on Suomi-NPP has been operating since 2011 and provides full global coverage at 375 meters resolutions [19]. Although satellite data has been broadly used in forest fire monitoring and in identifying the hotspots related to forest fire and biomass burning, the data is not available for all over a day. The schedule of satellite orbit limits data availability as the overpassing time is fixed once or twice a day. Besides, there is some delay in data processing overhead which cannot provide data in real-time.

The advancement in wireless sensor networks (WSN) has enabled the development of low-cost sensor networks which have been applied to many applications including environmental monitoring [20]. Distributed low-cost particulate matter (PM2.5) sensors were deployed in Xi'an, China to monitor air pollution [5]. Similar to the work in [2], 100 nodes of black carbon sensors were deployed across an urban area in West Oakland, California. The Outdoor Aerosol Sampler (OAS) system was proposed in [8]. The system used a low-cost PM2.5 sensor to measure the smoke concentration from a large prescribed fire event in Colorado, USA. In [17], the authors proposed a prototype of a low-cost sensor node to detect forest fire using mixed parameters (e.g., CO level, temperature, relative humidity, and light intensity). A common challenge of those forest fire monitoring systems is the network communication where commodity WiFi and cellular network are not always available. To overcome this limitation, the recent LoRaWAN protocol was applied for transmitting the sensor data from the remote forest area [16, 18]. With the low-cost, portable, and low power consumption benefits, the wireless sensor network could be used for large-scale deployment which can fulfill the gap in satellite data. However, most of the existing systems have been used in short-term observation or in laboratory testing where the robustness needs to be improved. In our case, the Canarin sensor node has been proven that it can operate all over a year.

In another important research related to the performance assessment of low-cost sensors, several works conducted experiments both in laboratory settings and ambient environments. In [1], the authors conducted extensive experiments to identify the potential of several candidates' PM2.5 sensors. The results revealed that the PM2.5 sensor using the light scattering technique correlated well with the standard equipment. An evaluation approach for the assessment of low-cost air quality sensors was proposed in [4]. The quality assessment of PM2.5 sensor in an ambient environment were also conducted in many research works [9, 11]. The study revealed that the high relative humidity (RH) factor significantly deviated from the PM2.5 reading. Similar to our work, we found that the accuracy of the forest detection model was significantly dropped in high-humidity weather.

6 CONCLUSIONS

In this paper, we share our first-hand experience in deploying a real-world field haze monitoring and forest fire detection network developed from a low-cost IoT sensor. The SEA-HAZEMON platform was developed to collect various air pollution parameters from the remote sensors. Unlike other platforms, our Canarin sensor nodes are self-sustain and robust where they can actively operate in the forest area for over a year. A simple model based on the concentrations of PM2.5 and CO was also applied for detecting real-time forest fire events. The accuracy of our forest fire detection system was evaluated through ground truth data collected from satellites and the forest fire control department during the peak forest fire season. Our model achieved an accuracy of more than 80%. However, we found that there were some false positive events reflecting the low precision and recall values. To address this issue, the relative humidity threshold is applied to remove faulty PM2.5 values caused by a high-humidity environment. Our study also revealed that the efficiency in detecting the fire hotspot located at higher altitudes is less. Besides, the effective area of the sensor node is also discovered which is less than 5km.

As for future work, we intend to deploy more sensor nodes in the study area which will improve the capability in detecting forest fire events. The study of plume movement affected by wind speed and direction is needed to be investigated. Furthermore, we plan to integrate the LoRa transmission on our sensor node which will support long-range communication.

ACKNOWLEDGMENTS

This research was supported by Asi@Connect program under grant contract ACA2016/376-562 and the Broadcasting and Telecommunications Research and Development Fund for Public Interest of Thailand (A63-1-2-005). We thank Pollution Control Department and Forest Fire Control Department of Thailand for their supports. We are also immensely grateful to Panjai Tantatsanawong, Veerachai Tanpipat and Doi Chang PaPae community for system deployment and maintenance. We would also like to show our gratitude to Prof. Kanchana Kanchanasut for creating this project and sharing her constructive advices with us.

REFERENCES

[1] David E Campbell Walter Ham Donald Schweizer Leland Tarnay Ahmed Mehadi, Hans Moosmüller and Julie Hunter. 2020. Laboratory and field evaluation of real-time and near real-time PM(2.5) smoke monitors. *Journal of the Air and Waste Management Association* 70, 2 (2020).

[2] Julien J. Caubel, Troy E. Cados, Chelsea V. Preble, and Thomas W. Kirchstetter. 2019. A Distributed Network of 100 Black Carbon Sensors for 100 Days of Air Quality Monitoring in West Oakland, California. *Environmental Science & Technology* 53, 13 (2019), 7564–7573.

[3] Janice L. Peterson John E. Core Paula Seamon Colin C. Hardy, Roger D. Ottmar. 2001. *SMOKE MANAGEMENT GUIDE FOR PRESCRIBED AND WILDLAND FIRE 2001 Edition.* National Wildfire Coordination Group. National Interagency Fire Center, Research Triangle Park, North Carolina, U.S.

[4] Barak Fishbain, Uri Lerner, Nuria Castell, Tom Cole-Hunter, Olalekan Popoola, David M. Broday, Tania Martinez Iñiguez, Mark Nieuwenhuijsen, Milena Jovasevic-Stojanovic, Dusan Topalovic, Roderic L. Jones, Karen S. Galea, Yael Etzion, Fadi Kizel, Yaela N. Golumbic, Ayelet Baram-Tsabari, Tamar Yacobi, Dana Drahler, Johanna A. Robinson, David Kocman, Milena Horvat, Vlasta Svecova, Alexander Arpaci, and Alena Bartonova. 2017. An evaluation tool kit of air quality micro-sensing units. *Science of The Total Environment* 575 (2017), 639–648.

[5] Meiling Gao, Junji Cao, and Edmund Seto. 2015. A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China. *Environmental Pollution* 199 (2015), 56–65.

[6] Gabriel Jobert, Pierre Barritault, Maryse Fournier, Salim Boutami, Daphnée Jobert, Adrien Marchant, Julien Michelot, Paul Monsinjon, Pierre Lienhard, and Sergio Nicoletti. 2020. Miniature particulate matter counter and analyzer based on lens-free imaging of light scattering signatures with a holed image sensor. *Sensors and Actuators Reports* 2, 1 (2020), 100010. https://www.sciencedirect.com/science/article/pii/S2666053920300072

[7] Thongchai Kanabkaew, Preechai Mekbungwan, Sunee Raksakietisak, and Kanchana Kanchanasut. 2019. Detection of PM2.5 plume movement from IoT ground level monitoring data. *Environmental Pollution* 252 (2019), 543–552. https://www.sciencedirect.com/science/article/pii/S0269749119306657

[8] S. Kelleher, C. Quinn, D. Miller-Lionberg, and J. Volckens. 2018. A low-cost particulate matter (PM2.5) monitor for wildland fire smoke. *Atmospheric Measurement Techniques* 11, 2 (2018), 1087–1097.

[9] Eemil Lagerspetz, Naser Hossein Motlagh, Martha Arbayani Zaidan, Pak L. Fung, Julien Mineraud, Samu Varjonen, Matti Siekkinen, Petteri Nurmi, Yutaka Matsumi, Sasu Tarkoma, and Tareq Hussein. 2019. MegaSense: Feasibility of Low-Cost Sensors for Pollution Hot-spot Detection. In *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, Vol. 1. 1083–1090. https://doi.org/10.1109/INDIN41052.2019.8971963

[10] William Lassman, Bonne Ford, Ryan W. Gan, Gabriele Pfister, Sheryl Magzamen, Emily V. Fischer, and Jeffrey R. Pierce. 2017. Spatial and temporal estimates of population exposure to wildfire smoke during the Washington state 2012 wildfire season using blended model, satellite, and in situ data. *GeoHealth* 1, 3 (2017), 106–121.

[11] Hai-Ying Liu, Philipp Schneider, Rolf Haugen, and Matthias Vogt. 2019. Performance Assessment of a Low-Cost PM2.5 Sensor for a near Four-Month Period in Oslo, Norway. *Atmosphere* 10, 2 (2019). https://www.mdpi.com/2073-4433/10/2/41

[12] Christopher O. Justice Louis Giglio, Wilfrid Schroeder. 2016. The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment* 178 (2016).

[13] NASA. 2022. Firms: Fire Information for Resource Management System. Retrieved August 18, 2022 from https://firms.modaps.eosdis.nasa.gov

[14] Office of Air Quality Planning and Standards Air Quality Assessment. July, 2018. *2014 National Emissions Inventory, version 2 Technical Support Document.* MIT Research Lab Technical Report. Division Emissions Inventory and Analysis Group, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, U.S.

[15] SEA-HAZEMON@TEIN project. 2021. *Final Report: Real-Time Haze Monitoring and Forest Fire Detection Information Centric Networks.* Technical Report. Asi@Connect. https://interlab.ait.ac.th/sea-hazemon-tein-project/report-final

[16] K Ram Prasanna, J.M. Mathana, T. Anne Ramya, and R. Nirmala. 2021. LoRa network based high performance forest fire detection system. *Materials Today: Proceedings* (2021). https://www.sciencedirect.com/science/article/pii/S2214785321042838

[17] Ridma Wanasinghe Kishanga Kottahachchi Udaya Dampage, Lumini Bandaranayake and Bathiya Jayasanka. 2022. Forest fire detection system using wireless sensor networks and machine learning. *Nature Scientific Reports* 12, 46 (2022).

[18] Roberto Vega-Rodríguez, Sandra Sendra, Jaime Lloret, Pablo Romero-Díaz, and Jose Luis Garcia-Navas. 2019. Low Cost LoRa based Network for Forest Fire Detection. In *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*. 177–184.

[19] Louis Giglio Ivan A. Csiszar Wilfrid Schroeder, Patricia Oliva. 2014. The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment* 143 (2014).

[20] Jennifer Yick, Biswanath Mukherjee, and Dipak Ghosal. 2008. Wireless sensor network survey. *Computer Networks* 52, 12 (2008), 2292–2330.