AUTOMATRIC AIR QUALITY MONITORING SYSTEM

Azhagar Vijay PS[1], Rohith MR[2], Kiruba Shankar P[3]
[1]UG Scholar, Electrical and Electronics Engineering& Bannari Amman Institute of Technology
[2] UG Scholar, Electrical and Electronics Engineering& Bannari Amman Institute of Technology
[3] UG Scholar, Electrical and Electronics Engineering& Bannari Amman Institute of Technology

Abstract - A technology known as an automatic air quality monitoring system continuously measures and evaluates the air quality in a specific environment. It is made to find and look at different pollutants and give real-time data for monitoring. These frameworks are regularly utilized in modern settings, urban communities, and homes to guarantee the air we inhale is protected and sound. For measuring the air quality, we are using ESP32 NODEMCU (microcontroller) with humidity and temperature sensors and a display unit for displaying the value of air quality. We are using IOT (Internet of Things) technology. and we are interfacing with cloud technology by accessing the data from the display and sending the message or notification to the mobile phone or tv and give the alarm sound. our main motto to fulfil the needs of industry people. By ensuring the level of air quality monitoring in their workplace. If the level of any gases like CO2 or some other gases increased means automatically the signal will be sent to the display system. By using this display module in their workplace, they can be priorly notified and take safety measures like wearing masks etc. so no one gets harmed by any type of diseases like lung diseases and cancer etc. This also impacts their health and reduces the chances of occurring any health issues by maintaining a moderate ambiance or as required. Hence, IoT-powered solutions provide better services to the industrialists, which in turn, provides better services to their customers. It creates a positive impact on human health by eliminating unwanted air. Pollutants and particulate matter by allowing the authorities to take decisions according to the situation.

***Key Words*:** Internet of Things, ESP32 NODEMCU, Detect gas level, Alarm sound.

1.INTRODUCTION

Over the last quarter of a century, companies developed rapidly. Such activities have created severe and complicated environmental issues [1]. Considering the importance of environmental quality in people's lives, the World Health Organization (WHO), by establishing limits to the amounts of different air contaminants, Ozone, nitrous oxides & Sulphur oxide which include ground level, has established recommendations on minimizing public and health consequences of air pollution [2][3]. The extreme climate is first & foremost contamination that has triggered climate erosion, climatic transition, stratosphere ozone depletion, habitat destruction, shifts in ecological and hydrological processes, soil degradation and pressures on Buildings for food processing, acid rain & global warming [4]. Occurrences of cancer, measles, asthma, respiratory problems, cardiovascular heart & chronic cardiovascular problems have been recorded for raising such pollutants. Therefore, the market for environmental emissions monitoring systems is growing [5] through sources of emissions utilizing harmful chemicals, these devices. will be able to identify and measure their origins easily. The modern air automated surveillance program uses laboratory analyzes with complicated facilities, large quantities, unreliable activities and high costs [6][7]. For large-scale construction, this renders high cost and wide volume difficult. This machine will only be built at essential control sites of some main firms,

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 10 | October - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

so device data cannot forecast the ultimate emissions situation. This thesis suggests integrating IoT technologies with environmental protection to resolve deficiencies in conventional control and detection approaches and to-research costs [8][9]. This work has been carried out based on many previous studies. In the past, studies performed air quality management and surveillance in the house [10]. This work is also focused on our study into remote contact for air quality monitoring. We also established an outdoor quality control program, in comparison to previous studies. A variety of substances like O3, SO2, CO and particulate matter can be calculated in the soil. Web sites track air quality remotely [11][12].

2. DATA QUALITY PRINCIPLES

It is common to find a definition of DQ from the consumer's point of view, where this trend is based on the treatment of data as a product. In Wang (1996), it is defined as "data that are fit for use by data consumers"; similar definitions are found in Barouch et al. (2016) and Liu et al. (2019). According to Barouch et al. (2016), the data consumer requires data to fulfill certain criteria that are essential for the tasks at hand. Being data a product, DQ is a multi-faceted concept since users have different expectations out of it. Thus, the DQ analysis has been divided into dimensions, where each dimension stands for an attribute that is important to the data consumer, or the application. After studying the term DQ in the field of IoT, we have identified several dimensions that can be relevant to the analysis of DQ. Table 1 present the most relevant DQ dimensions as well as their definitions and proposed evaluation metrics. In these tables, the first column is dimension name, and the second column includes a short definition of the dimension, which is a result of the review of several sources. It can be evidenced how a dimension can take several names, but it will have the same definition over different sources. Finally, the third column shows a formula or metric to

Table -1: DQ dimensions related to data values.

Dimension	Definition	Metric
Accuracy	"The degree to which data has attributes that correctly represent the true value of the intended attribute of a concept or event in a specific context of us" (Liu et al., 2019; ISO 25000 Portal, 2019).	$\alpha = \frac{\lvert v_m - v \rvert}{v}$, $DQ_{\text{accu}} = \max(0, 1 - \alpha)$, where v_m is the measured value, and v is the value accepted as true.
Precision	"The Precision is degree to which further measurements of the same phenomenon in a close time instant provides the same or similar results" (Sicari et al., 2018). It can be represented as the standard deviation of the measurement.	$DQ_{\text{prec}} = 1 - \frac{\sqrt{\frac{\sum_{i=1}^{n}(v_m - \bar{v}_m)^2}{n-1}}}{\bar{v}_m}$, where $\bar{v}_m$ is the mean of the measurement over n observations. The coefficient of variation is used to obtain a relative value.
Confidence	"The statistical error ε such that $[v - \varepsilon, v + \varepsilon]$ contains the real value with a confidence probability of p" (Klein & Lehner, 2009; Karkouch et al., 2016). It represents the statistical measurement error due to random environmental interference such as vibrations or shocks.	$\varepsilon = z \cdot \frac{\sigma}{\sqrt{n}}, n \geq 30$, $DQ_{\text{conf}} = 1 - \frac{\varepsilon}{\bar{v}_m}$, where z is the statistical value of the Z-distribution for a given confidence interval, while σ is the standard deviation of the population. Dividing by the mean give the relative margin of error.

3.DATA QUALITY IN THE CONTEXT OF AIR QUALITY ESTIMATION

In the context of air quality monitoring systems, The European Parliament and The Council has established Data Quality Objectives and Data Quality Indicators in the DIRECTIVE 2008/50/EC (UNION et al., 2008) guideline, while the Environmental Protection Agency (EPA) in the USA proposed the Quality Assurance Handbook for Air Pollution Measurement Systems (EPA, 2017) guideline. These documents define Data Quality Objectives (DQO) as the level of accepted threshold of the Data Quality Indicator (DQI), i.e., attributes of data quality. A close examination of these guidelines can lead to identify and match some of these indicators to the DQ dimensions previously discussed. We present below each DQI and its relationship with the DQ dimensions.

• Uncertainty: According to JCGM (2008), it is "a parameter associated with the result of a measurement that characterizes the dispersion of the values that could be reasonably attributed to the measurand." The authors also state that uncertainty is a generic term used to describe the sum of all sources of error associated with an environmental data operation. Uncertainty has two components, namely population uncertainty and measurement uncertainty. The former is related to the representativeness of the sample, while the latter is related to the precision, bias, and detection limit (EPA, 2017). Regarding the DQO for particulate matter pollutants, the maximum allowed uncertainty for fixed measurements (i.e., robust monitoring stations) is 25%, while for indicative measurements

(e.g., low-cost sensors measurements) is 50% (UNION et al., 2008). Based on this definition, this indicator is related to accuracy and confidence dimensions. • Minimum data capture: It has a limit of 90%, which means that the maximum number of missing values within one measurement period is 10% of the expected values (UNION et al., 2008). This indicator is related to completeness dimension.
• Minimum time coverage: This indicator for measurements of pollutants such as particulate matter (PM10/PM2.5) has a limit of 14% (1-day measurement per week at random, evenly distributed over the year, which would result on roughly 52 1-day measurements per year, or 8 weeks evenly distributed over the year, which would result on roughly 56 1-day measurements per year) (UNION et al., 2008). This indicator is related to timeliness and completeness dimensions.

4.SYSTEMATIC MAPPING METHOD

A systematic mapping study is a well-organized, and a frequently used methodology to synthesize the state of the art around a particular research area. This type of study looks for the "big picture" of some research topic, showing the branches and challenges associated with its James et al. (2016). This approach has been mainly used in software engineering; however, its application in the IoT field has been modest.

1.Research questions: In this step, the research questions are defined. These questions are expected to be solved when the systematic mapping process is completed.

2.Search strategy: This step defines the methodology of the research, starting by defining the "search chain" which will be applied to relevant academic databases.

3.Selection criteria: Inclusion and exclusion criteria are defined in this step. These criteria are used to filter the studies found in the previous step.

4. Data extraction: Once the Search Strategy and Selection Criteria are applied, relevant information about the Research Questions is extracted from the selected articles.

5. Analysis: In this step, we analyze the results obtained for drawing conclusion about the mapping study.

5.RESULT

This section presents the results of the mapping study developed to answer the research questions stated above. Before discussing the main results, we present a general overview of the papers under scope. The analysis of the main topics is presented around three aspects such as DQ dimensions and enhancement techniques, endangering factors, and DQ estimation and degradation. One of the first highlighting points is that the analysis of data quality in the context of IoT-based air quality monitoring systems is a topic with rising interest in the research community, especially in the last 7 years, with an average number of near 9 papers per year, as shown in Fig. 1. Even though there are some early approaches, such as by Harkat et al. (2006), the interest in DQ can be linked to the development and deployment of low-cost monitoring systems.

Figure 1: Histogram of paper publications in the context of data quality in IoT-based air quality monitoring systems per year.

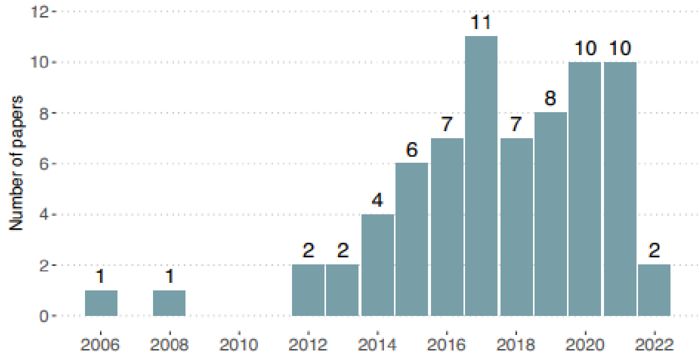

Figure 2: Illustrates the venues in which the analyzed works were published. Most of the papers (57.7%) were published in high quality journals (Q1 or Q2 according to the Simango ranking). Almost a third of the papers analyzed in this study (31%) were published in conferences.

.

Volume: 07 Issue: 10 | October - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

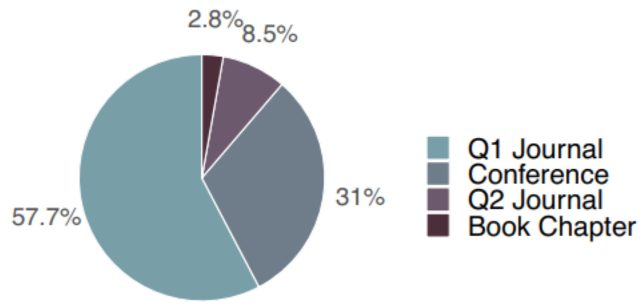

Figure 2: Venue of the publication

Regarding the variables of interest in the AQ monitoring system, Fig. 3 presents a histogram of the environmental variables identified in our study. The PM2.5 variable is the most frequently analyzed followed by the ozone (O3) and the nitrogen oxides (NOx). This result matches with the expectations, since low-cost PM and gas sensors are more prone to low-quality measurements as mentioned before.

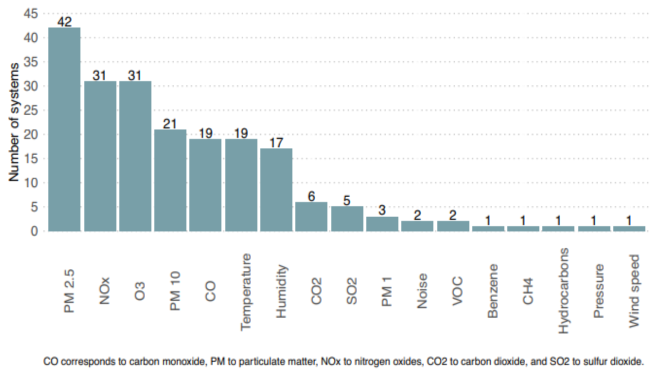

Fig. 3 Measured air-quality variables.

IoT-based air quality monitoring systems have been gaining popularity and are being included in a lot of new applications. Features like portability, small size, lightweight, low-cost, and first-hand data generation have motivated the creation of enthusiastic projects related to this topic. For these reasons, the trend shows that this approach will continue growing in the next decade. Figure 4 shows that most of the reviewed works (85.9%) are using low-cost sensors.

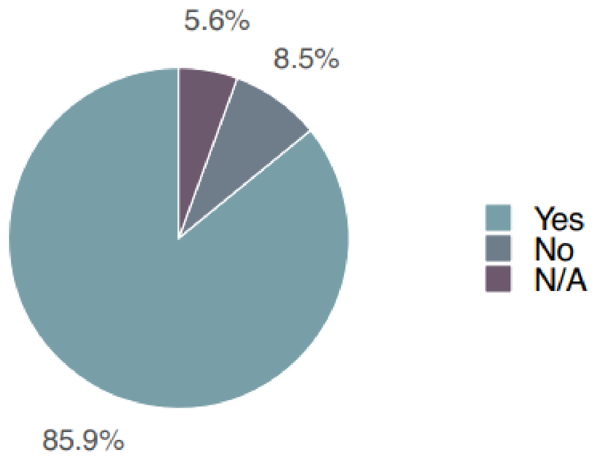

Fig. 4 Low-cost sensor usage

6.DISCUSSION

An increasing interest in analyzing the DQ topic is depicted in Fig. 1, which can be interpreted as the result of the number of low-cost IoT systems deployed for AQ monitoring (see Fig. 4). However, it was found in the reviewed papers that not many authors make an explicit mention of the DQ dimensions addressed in their work. What they do is to mention terms derived from "Data Quality," where DQ information is diffuse. We believe that the main reason for this phenomenon is the language of data quality, which has not been used in a proper formal way inside the IoT and air-quality monitoring applications yet. This is considered a serious issue to confront AQ measurement under the DQ definitions here presented. Accuracy was the most DQ indicator mentioned by different authors to measure "quality" inside AQ systems. Nevertheless, the introduction of other indicators will provide a more reliable and realistic approach inside the IoT AQ measurement. The minimum DQ dimensions and indicators that should be provided by a low-cost AQ system is a challenge that has to be established by different actors such as environmental agencies, enthusiastic developers, and technological industries around the world. We consider that environmental agencies have shown resistance in the implementation of portable and low-cost AQ supervision systems due to factors such as method of measurement, sensor aging, lack of redundancy, limited lifetime, and data error in storing/communication. Although these problems are serious and unresolved, low-cost AQ supervision will not be considered as a real alternative to determine the AQ in large-scale applications. On the other hand, low-cost sensors in the context of AQ applications have been growing as an alternative to empower citizens around the world. This tendency offers a lot of challenges and opportunities, which remarks the importance of adequate DQ definitions in those applications.

7. CONCLUSIONS

In this paper, we studied the data quality analysis on IoT-based air quality monitoring systems. First, we identified a general overview of data-quality dimensions within an IoT context. Then, data quality indicators and objectives in air-quality monitoring systems are reported, according to the guidelines by regulatory entities. Also, we propose mapping from indicators to dimensions to determine the relation between these concepts. In order to establish the state of data quality in IoT-based AQ systems, we developed a systematic mapping study about this field. The results showed an increasing number of studies that consider terms related to DQ within IoT-based air quality monitoring systems in the last few years; however, there is a lack of DQ terminology adoption and a rigorous application of DQ metrics. For instance, we had to identify the most relevant DQ dimensions related to IoT-based air quality system indirectly by analyzing the used enhancement techniques. To this end, we created a mapping between the enhancement techniques and the DQ dimensions. In general, we found authors do not use the terminology of the DQ field. We suppose this is due to two different factors. First, there is an absence of regulations that take into account indicative measurements (like low-cost sensor measurements) in the evaluation of air quality. Second, authors ignore the existing guidelines because they are not required to follow them. The primary objectives of their research are to evaluate technological alternatives or data processing techniques. It is understandable why low-cost sensor measurements are not fully considered by agencies in charge of environmental monitoring, because their data is prone to have more errors than a robust station. However, to avoid such distrust on low-cost sensors, an air quality monitoring system can be implemented to be DQ-aware and to include techniques to improve the quality of its data. In addition, many low-cost sensors can complement few robust stations to improve the resolution of the system, using the robust stations directly as references or sources of data to build reference models that help to improve DQ in low-cost air pollution sensors, for example, to be used in calibration processes.

8.REFERENCES

Alavi-Tottari, M., Williams, D., Salmond, J., & et al. (2013). Detection of malfunctions in sensor networks. Environmetrics,24(4),227–236. https://doi.org/10.1002/env.2206.

Alavi-Tottari, M., Salmond, J., Gaurnandanan, C., & et al. (2018). Automated data scanning for dense networks of low-cost air quality instruments: Detection and differentiation of instrumental error and local to regional scale environmental abnormalities. Environmental Modelling and Software,101,34–50. https://www.sciencedirect.com/science/article/abs/pii/S136481521631129X?via%3Dihub

Alvarado, M., Gonzalez, F., Fletcher, A., & et al. (2015). Towards the development of a low-cost airborne sensing system to monitor dust particles after blasting at open pit mine sites. Sensors (Switzerland),15(8),19,667–19,687. https://www.mdpi.com/1424-

Barcelo-Ordinals, J., Garcia-Vidal, J., Doudou, M., & et al. (2018). Calibrating low-cost air quality sensors using multiple arrays of sensors. pp. 1–6, https://ieeexplore.ieee.org/document/8377051

van Zeest, V., Osei, F., Stein, A., & et al. (2019). Calibration of low-cost NO2 sensors in an urban air quality network. Atmospheric Environment, 210, 66–75.
https://doi.org/10. 1016/j.atmosenv.2019.04.048.

Wang, L., Zhang, D., Yang, D., & et al. (2017). SPACETA: Cost-effective task allocation exploiting intradas and intermate correlations in sparse crowdsensing. ACM Transactions on Intelligent Systems and Technology, 9(2). https://doi.org/10.1145/3131671.

Wang, R. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5–34. https://doi.org/10.1080/07421222.1996.

Wang, W. C., Lung, S. C., Liu, C., & et al. (2020). Laboratory evaluations of correction equations with multiple choices for seed low-cost particle sensing devices in sensor networks. Sensors (Switzerland), 20(13), 1–25.
https://www.mdpi.com/1424-8220/20/13/3661