The Deployment of a LoRaWAN-Based IoT Air Quality Sensor Network for Public Good

James Montgomery Howerton
School of Data Science
University of Virginia
Charlottesville, VA
jh3df@virginia.edu

Benjamin Leo Schenck
School of Data Science
University of Virginia
Charlottesville, VA
bls7cc@virginia.edu

***Abstract*—The goals of this project are to implement a LoRaWAN-based network of air quality sensors in Charlottesville and to use its data to generate a comparative spatial model of air quality before and during the COVID-19 outbreak. The implementation of this network required the distribution of "The Things Network" (TTN) LoRa gateways and our own custom-made sensor kits to volunteers distributed throughout the city. Our sensor kits measure temperature, humidity, CO2, and Particulate Matter (PM) 2.5 and 10, allowing us to take measurements in line with the EPA's air quality index as well as to keep up with modern trends in research showing the importance of CO2 as an air quality metric. Preliminary spatial analysis comparing air quality before and after March 11, 2020, the day that UVA announced all classes would move online, shows a near universal decline in carbon dioxide levels, but inconclusive changes in particulate matter.**

***Index Terms*—Statistical Modeling, Environmental Systems**

I. Introduction

A foundational shift is taking place in environmental monitoring with the advancement of community-driven networks and low-cost, high-accuracy sensing technologies, creating an unprecedented opportunity to push forward the state-of-the-art in collaborative environmental research [1]. Traditional methods for ground-level monitoring have provided high accuracy, but low spatial resolution with high monetary costs. It has become clear in the past decade that the most direct route to distributed ground-level monitoring for higher spatial density involves community-led initiatives for environmental sensing. However, this approach comes with a few important caveats: established approaches have had very high accuracy standards and, in order to meet these standards, more sophisticated methods must be used to calibrate the chosen sensors and to correct for anomalies in their behavior. When the hardware, software, and infrastructure are publicly available, how should the data generated by this kind of system be governed and for what purposes should it be used? We have attempted to address these open questions by designing and implementing a sensor network as a collaboration between the School of Data Science, the Link Lab at the University of Virginia (UVA), and Smart Cville, a local nonprofit.

Our first step was to decide on a transport protocol for our sensor network. Our sponsors had already invested in TTN as a means to set up a LoRa network around Charlottesville, so we agreed to help expand their existing network. LoRa is a spread-spectrum modulation technique that allows low density data to be transmitted over long distances using low power [2], [3]. It also has the benefit of being a wireless technology that is suitable for low-energy applications without the need of a license. TTN is a community-based provider of LoRa technologies and infrastructures that can be used to collect and temporarily store transmitted data via the LoRa protocol. They are based on Free and Open Source technologies which better align with the interests and goals of our project. Over the course of our project, we have doubled the number of LoRa gateways in Charlottesville from 5 to 10, all housed in private homes around Charlottesville. We are the first members of the Charlottesville community and UVA to create a community-driven application for this network.

With the transport protocol in place, we designed an air quality (AQ) sensor kit. We have included functionality for sensing temperature, humidity, carbon dioxide (CO2), and particulate matter (PM) 2.5 and 10 variants. Our sensors utilize the LoRa network and are battery-powered. We have built and distributed 10 sensor kits around Charlottesville with help from a group of local volunteers. To provide open data access, we have utilized Grafana to visualize our temporal data and a mapping software provided by members of the TTN community to visualize our spatial data. Grafana is an Open Source analytics and monitoring solution that allows us to easily visualize time-series data from our sensors. Access to our online resources can be found in the linked appendix. To demonstrate the public utility of our sensor network, we have analyzed and visualized data from before and after March 11, the day that UVA announced that it would switch to online classes. Our visuals provide a detailed view of how emissions have changed throughout Charlottesville in this period. In what follows, we will discuss the literature on the topic of community-driven environmental monitoring, present our methods, data analytics, and conclusions with suggestions for future work.

II. Literature Review

As the world became more industrialized, high population density in urban centers and fossil fuel emissions resulted in rising levels of pollutants in the air. Research has shown that not only does polluted air affect respiratory health, but links have been found to increases in both atrial fibrillation [4] and ischemic stroke [5]. The World Health Organization (WHO) reports that 4.2 million deaths globally per year are attributed to outdoor air pollution alone and that over 90% of the population living in cities is exposed to particulate matter in concentrations exceeding the WHO guidelines [6]–[8]. In terms of economic damages, research has been able to quantify the negative effects of poor air quality in the hundreds of billions of dollars [9]. In addition to its impact on global warming, even moderate levels of CO2 have an impact on cognitive performance, thus presenting us with an important opportunity and responsibility to change our behavior [10].

Traditional approaches to monitoring air quality have been centralized by governmental agencies which uses highly expensive and accurate tools with low spatial density, generally one sensor per county [11]. One alternative to this method is satellite assessment which is able to achieve global coverage, but runs into several geospatial issues leading to low temporal resolution and problems taking measurements with cloud or vegetation cover [12]. To gain the advantages of both spatial and temporal resolution, smaller and cheaper ground sensors have become the new frontier in air quality sensing [13].

The "community science" approach has become more feasible than ever in recent years due to the release of low-cost and high-precision sensors which can consistently measure particulate matter among other pollutants. While these sensors should still be properly calibrated with reference instruments before having their accuracy assumed, their performance is a marked improvement over past generations, especially considering their price, size, and power consumption as attested by laboratory tests by independent researchers and the EPA [14], [15]. Along with these new sensors have come new proven transmission technologies such as LoRaWAN, enabling scientists and citizens to not just collect data but also reliably transport it to high performance computing infrastructures for processing [16].

Along with data collection, significant work has been done by independent researchers to determine proper techniques for calibrating low-cost sensors and detecting anomalies. Many of these approaches compare sensor readings from low-cost sensors to government-owned stations in the area, attempting to train models to calibrate sensors over time, space, or both [11], [17]. Several of these projects have shown success when using random forest models to calibrate their sensors and reinforce the improvements to accuracy possible when sensing multiple air quality metrics simultaneously [18]. Other projects report success in the application of LSTM Neural Network approaches to calibrate their sensors [19]–[21].

Many community-driven air quality sensing projects have sprouted up around the world utilizing these new technologies to implement real-world data collection and analysis schemes. Some examples include Air Quality Egg, Safecast Air, Publiclab, Smart Citizen, AirCasting, HackAIR, Luftdaten, PurpleAir, the Air Quality Data Commons, and more. These projects have pioneered but also encountered difficult challenges in community-led environmental sensing. Based on their experiences, we sought to expand their coverage by advancing the technical and social aspects of community-driven environmental monitoring in our community.

Potential use-cases for our environmental data are countless given the variety of sensors that are available today in addition to the number of researchers and volunteers around the world who are invested in this approach. Researchers have used wireless sensor networks to predict the spread of wildfires [22], [23], while others have focused on their use-cases in urban settings. Our contribution to Charlottesville has been in laying the groundwork for this sensor network so that local researchers and citizens may have the data they need to address air quality issues going forward.

III. Methodology

One of our major goals was to implement a networking infrastructure to facilitate the proliferation of community-driven, low-power sensor networks going forward. We have been successful in not only doubling the number of LoRa gateways throughout the city, but also in spreading knowledge about LoRa and its possibilities in the community and university, gathering ideas from multiple stakeholders in local organizations to spur collaborative projects with open technologies.

We began this process by holding events and meetings with both university and community members, where we explained details of the technology, presented similar projects around the world, and described the specifics of our project. We included members of UVA Schools of Data Science, Engineering, Architecture, Health, and the College of Arts & Sciences in our process of consultation. We also made an effort to include community members from local environmental groups C3 & LEAP, the Charlottesville city administration, and our sponsor Smart Cville. We facilitated discussion about the possibilities of this technology in our community, kept track of suggestions for future projects, and gathered a list of volunteers to host our prototype sensor kits. We filtered through the volunteers, choosing individuals who were located in major intersections and roads, but also distributed around the city to maximize our spread across Charlottesville. We gave both a LoRa gateway and a sensor kit to these individuals, however, future projects may no longer need to provide LoRa gateways along with their sensor kits once the city achieves sufficient network coverage.

Significant work went into designing and implementing all parts of our sensor kits, starting with the constraints that our sensor kits should take relatively high accuracy readings, have relatively long battery life, and have relatively low cost.

With these goals in mind, we began choosing electronic componentry. Our choice for the first prototype was to test the Adafruit Feather M0 with LoRa breakout (M0) as our microcontroller (MCU). This specific board allows us to supply both 5 volt (V) and 3.3V out while simultaneously enabling us to send LoRa packets easily using on-board hardware. We decided to use the Plantower PMS7003 PM sensor as it was the smallest, lowest-cost, and highest accuracy pre-calibrated sensor that we could find on the market. We chose to use the

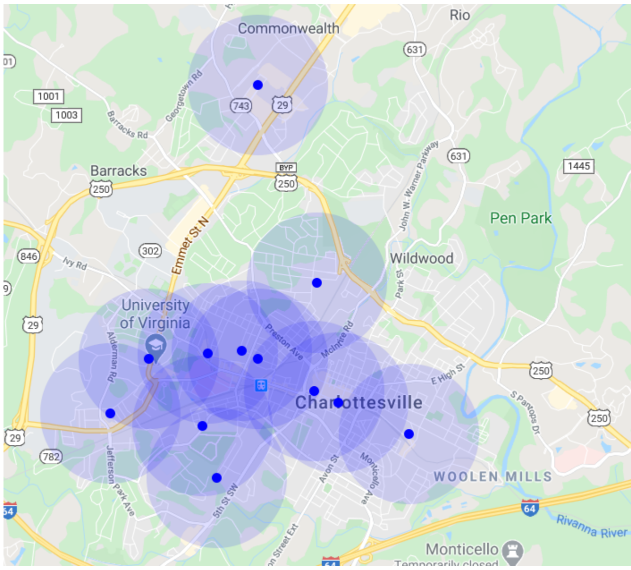

Fig. 1. LoRa Gateway Distribution (1km radius)

Sensirion SCD30 Nondispersive Infrared (NDIR) CO2 sensor considering its on-board temperature and humidity sensing, auto-calibration, and high accuracy [24].

We did not perform collocation tests for our sensors, but based our decisions on laboratory tests that have been conducted by the US Environmental Protection Agency as well as independent researchers who have shown that the PMS7003 factory calibration demonstrates low intramodel variability and strong correlation with reference instruments (a coefficient of variation less than 10% between units) [14], [15], [25]. Both PM and CO2 sensors are reported to be sensitive to environmental conditions, such as humidity and temperature, so we use these metrics to validate data points, but also report these metrics with our data set for future studies [26], [27].

Placement was also considered by our group, that is, how to better geographically situate our sensors. We sought to deploy our sensors in places of high urban traffic, so as to get an upper-bound on the air quality in the greater Charlottesville area. Once specific geographic locations were chosen, we ensured that our sensors were deployed in covered areas outdoors, would receive sunlight during some portion of the day to increase the lifespan of the device through energy harvesting, and were placed at a height between 1m and 10m to ensure that our readings reflected concentrations to which humans might be exposed.

Protecting our sensors from weather damage was of particular focus when designing the housing itself. For our pilot prototype, we modified plastic containers with 3D-printed mounts for our sensors with silicone coating (for the circuitry) and sealant. For our first hardware revision, we adopted a fully 3D-printed housing utilizing o-rings for weather sealing. Our pilot prototype used miniature breadboards with jumper wires to connect our components, however, we have designed a PCB to use in our first revision that simplified and improved the reliability of the setup greatly. The next steps in improving

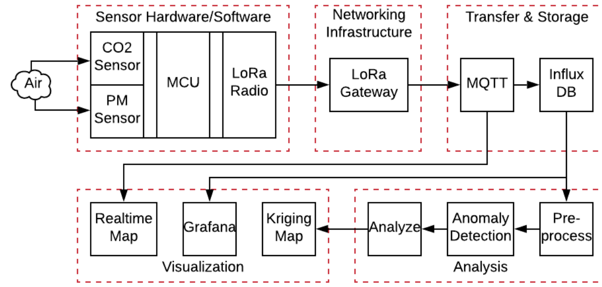

Fig. 2. Data Flow Diagram

the usability and reproducibility of our setup are provided in the future works section. Pictures of our assembly are provided in our digital appendix.

The software for this project consists of a single C script written in collaboration by us and our mentor. We utilize several community libraries to interface with our sensors and the radio module on our MCU. The Free and Open Source distribution of Arduino libraries significantly increased the speed with which we could prototype our design.

The revisions of both our hardware and software are available in our digital appendix. As illustrated in Figure 2, our data begins its life-cycle in the hardware sensors themselves, the PMS7003 and the SCD30. From there, it is formatted and transmitted to the embedded MCU. It is then transmitted via an on-board LoRaWAN radio module to a LoRa Gateway that receives the encoded packets, decodes them, and sends them to a Message Queuing Telemetry Transport (MQTT) broker. An instance of InfluxDB that is privately hosted on our server is subscribed to this MQTT broker and then pulls the data and stores it permanently. Our data visualization engine, Grafana, then pulls data from InfluxDB using a Secure Socket Layer (SSL) API call and updates our public time-series visualization dashboards. We also use an alternative visualization tool, made by the TTN community, that subscribes directly to our MQTT broker and generates a public real-time map of our data. To perform more complex analysis, we use a Python library to make SSL API calls to our InfluxDB instance, clean and aggregate our data, and then push the data to R for spatial analysis.

IV. Analysis

When beginning analysis on data from this kind of volunteer sensor network, the first thing to think about are the potential sources of noise in the dataset. We have sensors that only behave correctly under certain conditions, that are powered by unstable power sources, and that are not in certified enclosures. We must first remove erroneous data points caused by these issues before we move onto identifying truly anomalous events. We outlined realistic ranges for each of our metrics by looking at the data already gathered and reading up on the climate of Charlottesville. We then remove data points outside of these ranges and remove sets of data points that are

Metric	Minimum	Maximum
CO2 (ppm)	0	5000
PM 2.5 (ug/m^3)	0	350
PM 10 (ug/m^3)	0	350
Temperature (C)	0	41
Humidity (%)	0	100

Fig. 3. Acceptable Metric Bounds

perfectly linearly correlated in sequence, effectively removing long periods of constant values or linearly interpolated values. 24.43% of our original data points were removed through just these two simple data cleaning techniques. 49.58% of our readings had at least one missing value in them. As of April 15, 2020, we have recorded 46,309 sensor readings.

We then performed anomaly detection on the CO2, PM 2.5, Temperature, and Humidity features of our dataset using an isolation forest with an expected contamination of 1%. This was an exploratory analysis that led us to observe that the distribution of anomalies per-sensor kit was not uniform. We found that our 8th sensor kit was generating nearly twice as many anomalies as the next highest kit, spurring us to investigate the cause. We found that this sensor kit was placed very close to the railroad that runs through Charlottesville, indicating that the anomalies could be explained by trains periodically passing through. Sensor anomaly rates should approach their expected value over time due to the law of large numbers, meaning that this anomaly distribution will become more reliable the longer the sensors exist in the field. More figures detailing the results of our data cleaning and anomaly detection can be found in our digital appendix.

Plotting the correlation between the readings of our sensors for our three air quality metrics: CO2, PM 2.5, and PM 10 showed R^2 values of below 0.02. This showcases the massive amount of spatial density necessary to capture all of the air quality information in an urban setting. Further research must be done to determine the necessary sensor density to capture all of the meaningful air quality information in urban settings and, furthermore, to determine the most effective means to interpolate between sensor readings in scenarios with less than sufficient sensor coverage.

To generate a continuous map of air quality throughout the area in which we performed our study, we chose ordinary kriging as our interpolation method. Ordinary kriging is a form of Gaussian regression which gives weights to the values of the nearest spatial points (the sensors) and then interpolates to the entire raster. Using the formula $\hat{z}(x_0) = \sum_{i=1}^{N} \lambda_i z(x_i)$, where N is the number of sensors, $z(x_i)$ is the recorded value at the i^{th} sensor (mean of the recorded values over time in our case), and λ_i are the weights assigned to each sensor that are determined by geographic straight-line distance and spatial auto-correlation, $\hat{z}$ predictions are calculated for each pixel x_0 in the entire fine raster by summing across all the weighted terms. As this model is unbiased, the weights sum to one [28], [29]. From this, heatmaps can be visualized for the predictions

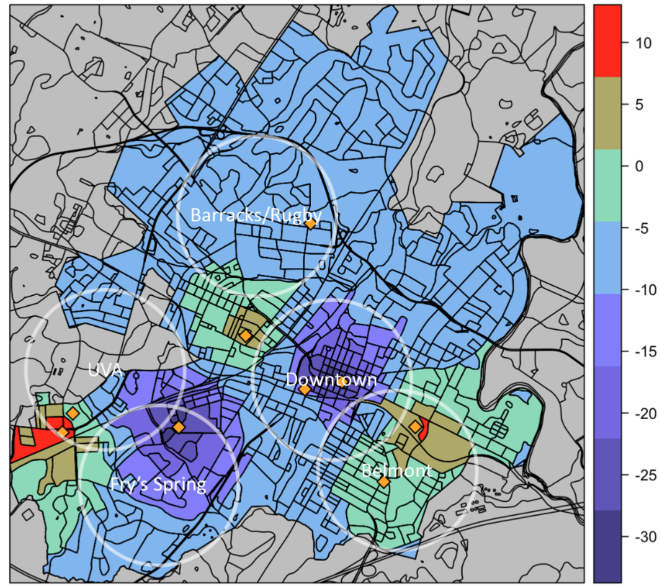

Fig. 4. Percent Change in Mean CO2 levels (ppm) since the March 11, 2020 announcement of the cancellation of in-person classes at UVA. The values shown on the map were interpolated using ordinary kriging. Orange diamonds represent sensor kits.

of each of the metrics. It must be noted however, that since kriging is a method of interpolation, and the coverage of sensors around the city is still fairly sparse, that these maps should not be currently used to inform Charlottesville residents of their air quality but, instead, be used as a proof-of-concept for when the sensors eventually cover a sizeable portion of the city in the future.

We were interested to see if we could use our network of sensors to see if there were changes in the air quality before and after March 11, 2020, the day that UVA cancelled in-person classes. We believe, given the UVA student body at the time made up about one third of the total Charlottesville population, that this day would be a reasonable estimate of the time that the greater Charlottesville community began "social distancing". As mentioned in the previous section, these heatmaps stem from collected data that is, at the moment, too sparse to make any confident statistical claims. However, as another proof-of-concept for when the network of sensors expands and more data becomes available, an example of the resultant heatmaps from kriging are shown in Fig. 3. While we are hopeful to have exponentially more data in the coming months and beyond, this map shows yet another potential application for disseminating meaningful information back to the community from the open data that has been collected.

V. Discussion and Future Work

Many of the decisions during the design and construction of our first prototype kits were made with relatively strict time and budget constraints. As such, our battery life is much lower than would be ideal, our first prototype kits are completely insulated, and our cost is higher than we would have liked for mass adoption. In the second phase of this project, work is already being done to solve all of these issues. A new dedicated MCU board is being designed with battery life as a top concern and new sensors are being evaluated for use. A fully 3D-printed enclosure is being re-designed to make our kits much easier and cheaper to build. These modifications will, in tandem, significantly reduce the cost of our sensor kits going forward, thus making them more accessible for community members and providing a better infrastructure from which to build future community sensing projects in Charlottesville and UVA.

Past projects in the community sensing space have used a variety of communication protocols to implement their data pipeline. Our experience with LoRa was relatively seamless, however, it was also expensive and did not perform as well in dense, urban areas. The LoRa gateways used for this project advertise a connection radius of 10km with line-of-sight, but we experienced a radius closer to 1 km, reflected in the radius chosen for Figure 1. TTN gateways cost around $200 dollars each, forcing us to rely on outside funding from the UVA Link Lab in order to increase our network coverage. Other projects at UVA have had significant difficulty in working with LoRa devices that are not 'plug-and-play' as the TTN ones are. With all of this taken into account, we can say confidently that LoRa is the best low-power wireless sensing networking technology that we have used, but that future researchers should evaluate alternatives that can solve these issues of high expense and over-promised connection radius.

While most data scientists rely on the Open Source community for their ongoing support of Python and R, our project has attempted to go all-in on the 'open' approach. The software running on our boards, the hardware designs and wiring diagrams, the Jupyter Notebooks that we use to analyze our data, and the data itself, among other resources, are all available using open licenses. TTN is an open network, the MCU that we chose has open schematics, even the company that we have chosen to print our PCBs, OSH Park, is a proponent of the Open Hardware movement. In our experience, we see the Open Source software, hardware, and data communities as a perfect fit for projects existing at the crossroads between the academic and the 'hacker' communities in our university and the city that surrounds it.

The goals set forth at the beginning of our project were to increase the size of the LoRa network in Charlottesville, design and distribute sensor kits that utilize that network, provide open access to our collected data, and demonstrate the utility of our data through analysis. We were successful in all of our initial goals, but there is still much work to be done. There is more insight that can be gained as the quantity of sensors and collected data increases. One visualization that can be created from this is an animated collection of heatmaps. Similar to how weather stations have historical and future animated weather radars, showing trends in the air quality with respect to time can be a way for Charlottesville residents to intuitively get a sense of the relative air quality at a given time.

From a methodological perspective, implementing an array of anomaly detection methods for our raw data might help improve the accuracy of correctly identified data points. While we used the isolation forest method to detect anomalies in our data, there are a variety of alternative methods that may result in better performance, especially in aggregate. Of particular interest are more recent developments in neural network architectures, such as LSTMs [21].

Healthcare is another potential field where air quality data can be of great use. We have received Institutional Review Board approval to have access to the health records of all respiratory illness admits to the UVA Health System starting in February 2020. We will also be granted access to the patient home locations (general census tracts to protect privacy), and date of admission. Future work can utilize this data for a time-series analysis, determining if there are any density-adjusted correlations between air quality and the number of hospital admissions based on location within the city, akin to other related research projects from around the world [30], [31].

With the unprecedented actions taken by the local, state, and federal government regarding the COVID-19 pandemic, analysis of the air quality data before and after the implementation of "social distancing", as well as after the crisis comes to an effective conclusion could be enlightening. By analyzing which areas of the city show significant decreases in air pollutants after the majority of the population stays mostly housebound, it may provide insight as to which areas the Charlottesville policymakers should focus their efforts on combating the harmful human effects on the quality of the air. Similarly, if there are air pollutant spikes after "social distancing" has concluded, it may provide even more evidence for the locations of highest human influence. Experimental research can also be advanced on the role that PM plays in the dispersion of the virus, as it is known in the environmental science community that PM is a carrier for both inorganic and organic matter [32]–[34].

We are hopeful that, in addition to further expanding our infrastructure across the city of Charlottesville, our research will inspire the inception and continuation of other citizen science air quality sensor projects across the state, country, and world. As we discovered during the course of our project, the government mandated air quality sensing is sparse both spatially and temporally, and furthermore, not easily obtainable by the public. Expanding the reach of a network of air quality sensors fueled by open data can bring huge benefits in terms of environmental public engagement, especially in the age of climate change.

Acknowledgments

We thank our mentors and collaborators Luis Felipe Rosado Murillo, Jonathan Kropko, Lucas Ames, Jon Goodall, and Adam Slez for ensuring the success of this project.

Appendix

Our appendix can be accessed online at:
thejimster82.github.io/CvilleAQ

REFERENCES

[1] E. G. Snyder, T. H. Watkins, P. A. Solomon, E. D. Thoma, R. W. Williams, G. S. W. Hagler, D. Shelow, D. A. Hindin, V. J. Kilaru, and P. W. Preuss, "The Changing Paradigm of Air Pollution Monitoring," *Environmental Science & Technology*, vol. 47, no. 20, pp. 11 369–11 377, Oct. 2013, publisher: American Chemical Society.

[2] J. Haxhibeqiri, E. D. Poorter, I. Moerman, and J. Hoebeke, "A Survey of LoRaWAN for IoT: From Technology to Application," *Sensors*, vol. 18, no. 11, p. 3995, Nov. 2018, number: 11 Publisher: MDPI.

[3] I. T. Ali and R. F. Sari, "Research Opportunities of LoRaWAN for Internet of Things Implementation," in *2018 2nd International Conference on Applied Electromagnetic Technology (AEMT)*, Apr. 2018, pp. 61–66.

[4] A. G. Solimini and M. Renzi, "Association between Air Pollution and Emergency Room Visits for Atrial Fibrillation," *International Journal of Environmental Research and Public Health*, vol. 14, no. 6, Jun. 2017.

[5] J.-W. Chung, O. Y. Bang, K. Ahn, S.-S. Park, T. H. Park, J. G. Kim, Y. Ko, S. Lee, K. B. Lee, J. Lee, K. Kang, J.-M. Park, Y.-J. Cho, K.-S. Hong, H.-W. Nah, D.-H. Kim, J.-K. Cha, W.-S. Ryu, D.-E. Kim, J.-T. Kim, J. C. Choi, M.-S. Oh, K.-H. Yu, B.-C. Lee, J. S. Lee, J. Lee, H.-K. Park, B. J. Kim, M.-K. Han, and H.-J. Bae, "Air Pollution Is Associated With Ischemic Stroke via Cardiogenic Embolism," *Stroke*, vol. 48, no. 1, pp. 17–23, 2017.

[6] "WHO | Concentrations of fine particulate matter (PM2.5)," library Catalog: www.who.int Publisher: World Health Organization.

[7] "WHO | WHO Global Ambient Air Quality Database (update 2018)," library Catalog: www.who.int Publisher: World Health Organization.

[8] "9 out of 10 people worldwide breathe polluted air, but more countries are taking action," library Catalog: www.who.int.

[9] A. L. Goodkind, C. W. Tessum, J. S. Coggins, J. D. Hill, and J. D. Marshall, "Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions," *Proceedings of the National Academy of Sciences*, vol. 116, no. 18, pp. 8775–8780, Apr. 2019.

[10] Satish Usha, Mendell Mark J., Shekhar Krishnamurthy, Hotchi Toshifumi, Sullivan Douglas, Streufert Siegfried, and Fisk William J., "Is CO2 an Indoor Pollutant? Direct Effects of Low-to-Moderate CO2 Concentrations on Human Decision-Making Performance," *Environmental Health Perspectives*, vol. 120, no. 12, pp. 1671–1677, Dec. 2012.

[11] A. L. Clements, W. G. Griswold, A. Rs, J. E. Johnston, M. M. Herting, J. Thorson, A. Collier-Oxandale, and M. Hannigan, "Low-Cost Air Quality Monitoring Tools: From Research to Practice (A Workshop Summary)," *Sensors*, vol. 17, no. 11, p. 2478, Nov. 2017, number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

[12] P. Gupta, P. Doraiswamy, R. C. Levy, O. Pikelnaya, B. Feenstra, J. Maibach, A. Polidori, F. Kiros, and K. Mills, "Integration of satellite observations with low-cost air quality monitors: A data science approach to monitor air quality," *AGU Fall Meeting Abstracts*, vol. 43, Dec. 2018.

[13] S. J. Johnston, P. J. Basford, F. M. J. Bulot, M. Apetroaie-Cristea, N. H. C. Easton, C. Davenport, G. L. Foster, M. Loxham, A. K. R. Morris, and S. J. Cox, "City Scale Particulate Matter Monitoring Using LoRaWAN Based Air Quality IoT Devices," *Sensors*, vol. 19, no. 1, p. 209, Jan. 2019, number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[14] M. Badura, P. Batog, A. Drzeniecka-Osiadacz, and P. Modzel, "Regression methods in the calibration of low-cost sensors for ambient particulate matter measurements," *SN Applied Sciences*, vol. 1, no. 6, p. 622, May 2019.

[15] A. S. Ali, H. Zhao, and B. Stephens, "An evaluation of low-cost particulate matter and CO2 sensors and their applications," p. 2, 2019.

[16] L. Parri, S. Parrino, G. Peruzzi, and A. Pozzebon, "Low Power Wide Area Networks (LPWAN) at Sea: Performance Analysis of Offshore Data Transmission by Means of LoRaWAN Connectivity for Marine Monitoring Applications," *Sensors (Basel, Switzerland)*, vol. 19, no. 14, Jul. 2019.

[17] P.-Y. Yin, C.-C. Tsai, R.-F. Day, C.-Y. Tung, B. Bhanu, 1 Department of Information Management, National Chi Nan University, Nantou, 54561, Taiwan, ROC, 2 Institute of Strategy and Development of Emerging Industry, National Chi Nan University, Nantou, 54561, Taiwan, ROC, and 3 Center for Research in Intelligent Systems, University of California, Riverside, California 92521, USA, "Ensemble learning of model hyperparameters and spatiotemporal data for calibration of low-cost $PM_{2.5}$ sensors," *Mathematical Biosciences and Engineering*, vol. 16, no. 6, pp. 6858–6873, 2019.

[18] N. Zimmerman, A. A. Presto, S. P. N. Kumar, J. Gu, A. Hauryliuk, E. S. Robinson, A. L. Robinson, and R. Subramanian, "A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring," *Atmospheric Measurement Techniques*, vol. 11, pp. 291–313, Jan. 2018.

[19] S. Maya, K. Ueno, and T. Nishikawa, "dLSTM: a new approach for anomaly detection using deep learning with delayed prediction," *International Journal of Data Science and Analytics*, vol. 8, no. 2, pp. 137–164, Sep. 2019.

[20] E. U. Warriach and K. Tei, "A comparative analysis of machine learning algorithms for faults detection in wireless sensor networks," *International Journal of Sensor Networks*, vol. 24, no. 1, pp. 1–13, Jan. 2017.

[21] A. Verner, "LSTM Networks for Detection and Classification of Anomalies in Raw Sensor Data," Ph.D. dissertation, Apr. 2019.

[22] J. Lloret, M. Garcia, D. Bri, and S. Sendra, "A Wireless Sensor Network Deployment for Rural and Forest Fire Detection and Verification," *Sensors*, vol. 9, no. 11, pp. 8722–8747, Nov. 2009.

[23] H. Soliman, K. Sudan, and A. Mishra, "A smart forest-fire early detection sensory system: Another approach of utilizing wireless sensor and neural networks," in *2010 IEEE SENSORS*, Nov. 2010, pp. 1900–1904.

[24] Y. Liu, X. Ni, Y. Wu, and W. Zhang, "Study on effect of temperature and humidity on the CO 2 concentration measurement," *IOP Conference Series: Earth and Environmental Science*, vol. 81, p. 012083, Aug. 2017.

[25] F. M. J. Bulot, S. J. Johnston, P. J. Basford, N. H. C. Easton, M. Apetroaie-Cristea, G. L. Foster, A. K. R. Morris, S. J. Cox, and M. Loxham, "Long-term field comparison of multiple low-cost particulate matter sensors in an outdoor urban environment," *Scientific Reports*, vol. 9, no. 1, pp. 1–13, May 2019, number: 1 Publisher: Nature Publishing Group.

[26] S. S. Shrestha and G. M. Maxwell, "An experimental evaluation of HVAC-Grade carbon dioxide sensors–part I: test and evaluation procedure," *ASHRAE Transactions*, vol. 115, no. 2, pp. 471–484, Jul. 2009, publisher: American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc. (ASHRAE).

[27] H.-Y. Liu, P. Schneider, R. Haugen, and M. Vogt, "Performance Assessment of a Low-Cost PM2.5 Sensor for a near Four-Month Period in Oslo, Norway," *Atmosphere*, vol. 10, no. 2, p. 41, Feb. 2019.

[28] D. G. Krige, "A statistical approach to some basic mine valuation problems on the Witwatersrand," *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 52, no. 6, pp. 119–139, Dec. 1951, publisher: Southern African Institute of Mining and Metallurgy.

[29] "Kriging | Columbia University Mailman School of Public Health."

[30] A. K. Gorai, F. Tuluri, and P. B. Tchounwou, "A GIS Based Approach for Assessing the Association between Air Pollution and Asthma in New York State, USA," *International Journal of Environmental Research and Public Health*, vol. 11, no. 5, pp. 4845–4869, May 2014.

[31] W. Wang, Y. Ying, Q. Wu, H. Zhang, D. Ma, and W. Xiao, "A GIS-based spatial correlation analysis for ambient air pollution and AECOPD hospitalizations in Jinan, China," *Respiratory Medicine*, vol. 109, no. 3, pp. 372–378, Mar. 2015.

[32] J. W. Park, Y. H. Lim, S. Y. Kyung, C. H. An, S. P. Lee, and S. H. Jeong, "Detection of Pathogenic Viruses in the Atmosphere during Asian Dust Events in Incheon City," *Tuberculosis and Respiratory Diseases*, vol. 59, no. 3, pp. 279–285, Sep. 2005.

[33] J. Ciencewicki and I. Jaspers, "Air Pollution and Respiratory Viral Infection," *Inhalation Toxicology*, vol. 19, no. 14, pp. 1135–1146, Jan. 2007, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/08958370701665434.

[34] L. Setti, F. Passarini, G. de Gennaro, A. Di Gilio, J. Palmisani, P. Buono, G. Fornari, M. G. Perrone, A. Piazzalunga, P. Barbieri, E. Rizzo, and A. Miani, "Evaluation of the potential relationship between Particulate Matter (PM) pollution and COVID-19 infection spread in Italy," *The Italian Society of Environmental Medicine*, 2020.