



International Journal of Marketing Management

ISSN 2454 - 5007



www.ijmm.net

Email ID: editor@ijmm.net , ijmm.editor9@gmail.com

Digital Twin of Atmospheric Environment: Sensory Data Fusion for High-Resolution PM2.5 Estimation and Action Policies Recommendation

Md Khaja Pasha [1], Abdul Rahman [2], Faizaan Ilyas [3],

Saif Ali [4], Guntuka Nitish [5],¹Assistant professor,

^{2,3,4,5}Students, Department of Electronics and Communication Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana.

Abstract

This study addresses the health risks posed by particulate matter smaller than 2.5 microns (PM2.5), which can penetrate deep into the lungs and bloodstream, leading to respiratory and cardiovascular issues. To overcome the inefficiencies and limited coverage of traditional monitoring methods, a digital twin (DT) of the atmospheric environment is developed, integrating remote sensing and observational data with feedback mechanisms for dynamic representation. PM2.5 levels are estimated using an ensemble of Random Forest and Gradient Boosting models, yielding robust predictions with high accuracy (RMSE of 38.94, R^2 of 0.728 within a 95% confidence interval). The predicted values inform actionable recommendations such as cloud seeding to disperse pollutants, scheduled air filtering to reduce indoor pollution, and public SMS alerts to encourage preventive measures during high pollution days. An optimization problem is formulated to identify the most impactful action policies. Additionally, the study quantitatively assesses variable contributions to PM2.5 levels, demonstrating the potential of digital twins in enhancing environmental monitoring and management while offering strategies to mitigate the adverse health effects of air pollution.

I. Introduction

Air pollution is one of the most pressing environmental challenges, causing significant health risks worldwide. According to the World Health Organization (WHO), air pollution results in 4.2 million deaths annually, and 9 out of 10 people breathe air that exceeds WHO guideline limits. Among the various pollutants, PM2.5—particulate matter smaller than 2.5 micrometers—is particularly hazardous due to its ability to penetrate deep into the lungs, causing severe respiratory and cardiovascular diseases. Developing effective policies for PM2.5 mitigation is vital, yet traditional ground monitoring systems are sparse, costly, and limited in coverage, making them inadequate for timely preventive measures in many regions.

Advances in remote sensing and meteorological data modeling provide promising alternatives for PM2.5 forecasting. Early models, such as Multiple Linear Regression (MLR), Mixed-Effect Model (MEM), Chemical Transport Model (CTM), and Geographically Weighted Regression (GWR), have been explored

extensively. Each model has specific strengths and weaknesses based on its application scenario. For example, MEM performs well with scarce data, CTM is better for global-scale predictions, and GWR excels in local settings, while MLR has shown the weakest predictive capabilities. Incorporating auxiliary meteorological variables and land-use data has been found crucial in enhancing model accuracy and performance.

The emergence of more sophisticated machine learning (ML) and deep learning (DL) techniques has shifted the focus of PM_{2.5} estimation methods. However, practical applications and integration of these models into broader frameworks remain underexplored. The concept of Digital Twins (DT)—digital replicas of environments or objects with bidirectional information exchange—has gained traction as a potential solution. In air quality control, a DT pipeline can utilize model predictions to make informed and automated decisions, such as preventive actions. Despite its promise, critical questions remain open regarding the implementation of DTs for PM_{2.5} management, paving the way for future research and development in this field

II. Literature Review:

These references encompass diverse studies in environmental health, advanced modeling techniques, and innovative technologies. Hoek et al. (2013) reviewed long-term air pollution exposure and its link to cardio-respiratory mortality, highlighting the variability in effects due to population and methodological differences. Brook et al. (2010) examined the causal relationship between particulate matter exposure and cardiovascular diseases. Lundberg and Lee (2017) introduced SHAP for interpreting model predictions, offering consistency in feature importance. Grieves and Vickers (2017) explored Digital Twins as virtual replicas for predicting complex system behavior. Pylianidis et al. (2021) discussed Digital Twin applications in agriculture to enhance system control. Ketzler et al. (2020) reviewed the role of Digital Twins in urban planning for city management improvements. Tobías et al. (2020) analyzed air quality changes during the COVID-19 lockdown in Barcelona, noting pollutant reductions and ozone increases. Lastly, Lim et al. (2021) presented the Temporal Fusion Transformer for interpretable and high-performing multi-horizon time series forecasting, applicable across datasets. This collection reflects advancements in environmental monitoring, predictive analytics, and decision-making frameworks.

III. Digital Twin of the Atmospheric Environment:

A Digital Twin of the Atmospheric Environment integrates sensory data fusion and predictive modeling to enhance the estimation of PM_{2.5} concentrations and provide actionable policy recommendations. It represents a dynamic virtual replica of the physical atmospheric system, combining remote sensing and ground-

based observational data with feedback mechanisms for real-time updates. Advanced machine learning models, such as Random Forest and Gradient Boosting, are employed for high-resolution PM2.5 predictions, with optimization frameworks to identify impactful mitigation strategies. These strategies include measures like cloud seeding, air filtering schedules, and public notifications for high-pollution days. By incorporating sensory data, real-time interaction, and iterative improvements, the Digital Twin creates a robust platform for monitoring and managing air quality, offering innovative solutions to reduce pollution-related health risks.

Raspberry Pi in Estimation of PM2.5:

The A Digital Twin of the Atmospheric Environment can leverage the following components to create a robust system for PM2.5 estimation and policy recommendation. The hardware setup includes a Raspberry Pi Pico Microcontroller connected to gas sensors (MQ2, MQ7, MQ135) and a PM2.5 Dust Smoke Particle Sensor to monitor air quality parameters. An OLED display provides real-time visual feedback, while a WiFi module enables wireless communication for data transmission. A GSM module can be integrated to send alerts, such as SMS notifications, for preventive measures during high pollution levels. The system also includes an air valve motor for controlling air flow, and a buzzer for alarms or warnings. To support these functionalities, the software components include programming in Embedded C and Micro Python, alongside the Arduino IDE for development and deployment. This combination facilitates sensory data fusion, real-time interaction, and actionable recommendations, aligning with the dynamic and high-resolution capabilities required in the Digital Twin framework.

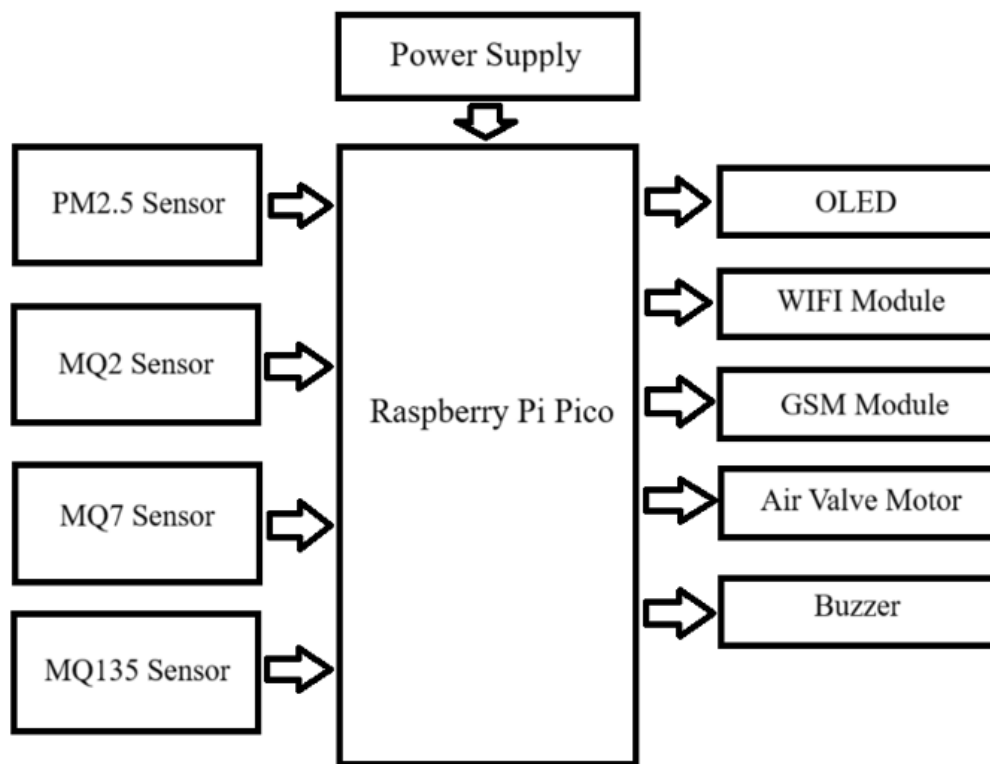
IV. Proposed System

The proposed Digital Twin (DT)-based PM2.5 system integrates advanced technologies and data fusion techniques to monitor and mitigate air pollution effectively. It begins with the collection and fusion of remote sensing data, such as satellite imagery, and ground observational data from air quality monitors to create a comprehensive dataset. Machine learning models like Random Forest and Gradient Boosting are employed to estimate PM2.5 levels, enhanced by feedback mechanisms for adjusting future input data. The trained DT model provides high-resolution PM2.5 predictions and analyzes key variables influencing pollution. An optimization framework is used to identify impactful action policies, such as cloud seeding to disperse pollutants, air filtering schedules to minimize indoor pollution, and SMS notifications to inform the public about high pollution days.

The system advances through implementation and continuous monitoring in real-world scenarios, ensuring the effectiveness of the recommended policies. Real-time data and feedback loops refine the DT model to improve

future predictions and recommendations, while metrics such as RMSE and R^2 are regularly evaluated to assess model performance. Efforts to incorporate new data sources and algorithms aim to enhance the system's accuracy and reliability further. This holistic approach highlights the potential of DT technology in addressing PM2.5 pollution challenges and offers a dynamic solution for protecting human health and improving air quality across regions.

V. Block Diagram



Block diagram of Digital Twin of Atmospheric Environment

Raspberry Pi Pico Microcontroller:

The Raspberry Pi Pico is a microcontroller board based on the Raspberry Pi RP2040 microcontroller chip. Like Raspberry Pi computers, Raspberry Pi Pico features a pin header with 40 connections, along with a new debug connection enabling you to analyze your programs directly from another computer (typically by connecting it directly to the GPIO pins on a Raspberry Pi). Raspberry Pi Pico is a brand new, low-cost, yet highly flexible development board designed around a custom-built RP2040 microcontroller chip designed by Raspberry Pi. Raspberry Pi Pico – ‘Pico’ for short – features a dual-core Cortex-M0+ processor (the most energy-efficient Arm processor available), 264kb of SRAM, 2MB of flash storage, USB 1.1 with device and host support, and a wide range of flexible I/O options.



Fig 4.7: Raspberry Pi Pico

Features of Raspberry Pi Pico:

- 21 mm × 51 mm form factor
- RP2040 microcontroller chip designed by Raspberry Pi in the UK
- Dual-core Arm Cortex-M0+ processor, flexible clock running up to 133 MHz
- 264KB on-chip SRAM

- 2MB on-board QSPI Flash
- 26 multifunction GPIO pins, including 3 analog inputs
- 2 × UART, 2 × SPI controllers, 2 × I2C controllers, 16 × PWM channels
- 1 × USB 1.1 controller and PHY, with host and device support
- 8 × Programmable I/O (PIO) state machines for custom peripheral support
- Supported input power 1.8–5.5V DC
- Operating temperature -20°C to +85°C
- Castellated module allows soldering direct to carrier boards
- Drag-and-drop programming using mass storage over USB
- Low-power sleep and dormant modes
- Accurate on-chip clock
- Temperature sensor
- Accelerated integer and floating-point libraries on-chip

Raspberry Pi Pico Pin Description:

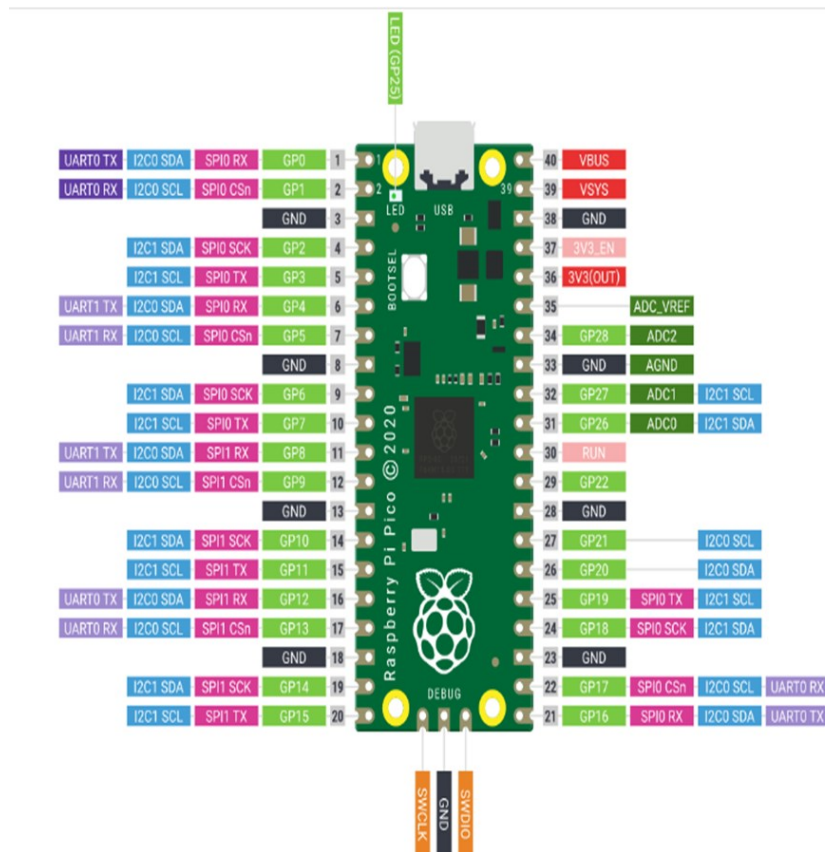


Fig- Raspberry Pi Pico Pinout

GAS SENSOR

A gas detector is a device which detects the presence of various gases within an area, usually as part of a safety system. This type of equipment is used to detect a gas leak and interface with a control system so a process can be automatically shut down. A gas detector can also sound an alarm to operators in the area where the leak is occurring, giving them the opportunity to leave the area. This type of device is important because there are many gases that can be harmful to organic life, such as humans or animals.



Fig -: Gas Sensor

FEATURES

- High sensitivity to LPG, natural gas, town gas.
- Small sensitivity to alcohol, smoke.
- Fast response.
- Stable and long life.
- Simple drive circuit.

OLED (Organic Light Emitting Diodes):

OLED (Organic Light Emitting Diodes) is a flat light emitting technology, made by placing a series of organic thin films between two conductors. When electrical current is applied, a bright light is emitted. OLEDs are emissive displays that do not require a backlight and so are thinner and more efficient than LCD displays (which do require a white backlight).

OLED displays are not just thin and efficient - they provide the best image quality ever and they can also be made transparent, flexible, foldable and even rollable and stretchable in the future. OLEDs represent the future of display technology.

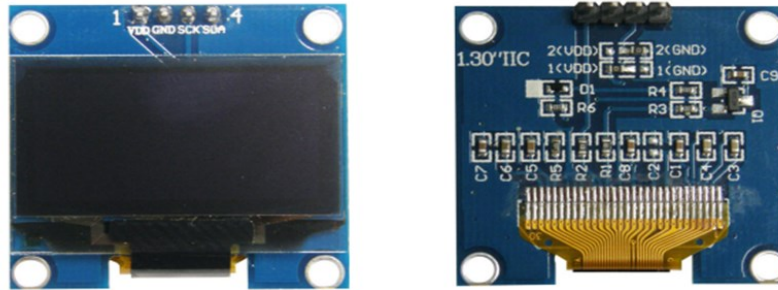


Fig: OLED Module

PM2.5 Sensor Module

PM2.5 Dust Smoke Sensor Module GP2Y1010AU0F is an optical air quality sensor, designed to sense dust particles. It is capable of detecting fine particle larger than $0.8\mu\text{m}$ in diameter, even like the cigarette smoke. An infrared emitting diode and a phototransistor are diagonally arranged into this device, to allow it to detect the reflected light of dust in air. It is especially effective in detecting very fine particles like cigarette smoke, and is commonly used in air purifier systems. Analog voltage output of the sensor is linear with dust density. The module has embedded voltage boost circuit to support wide range of power supply.

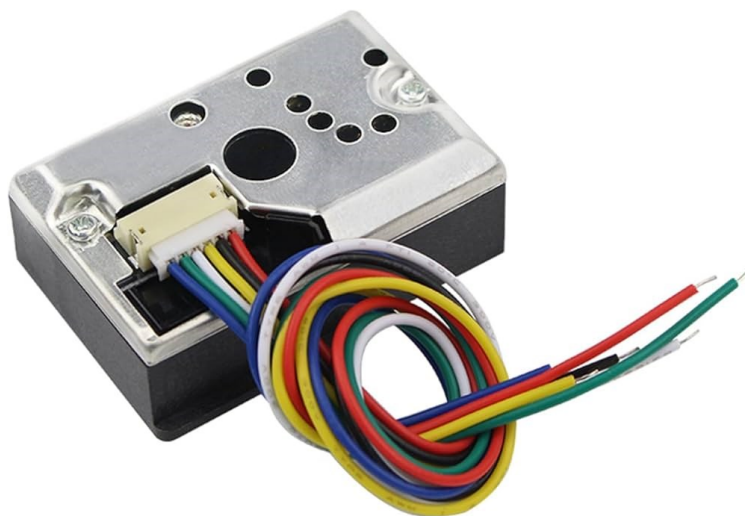
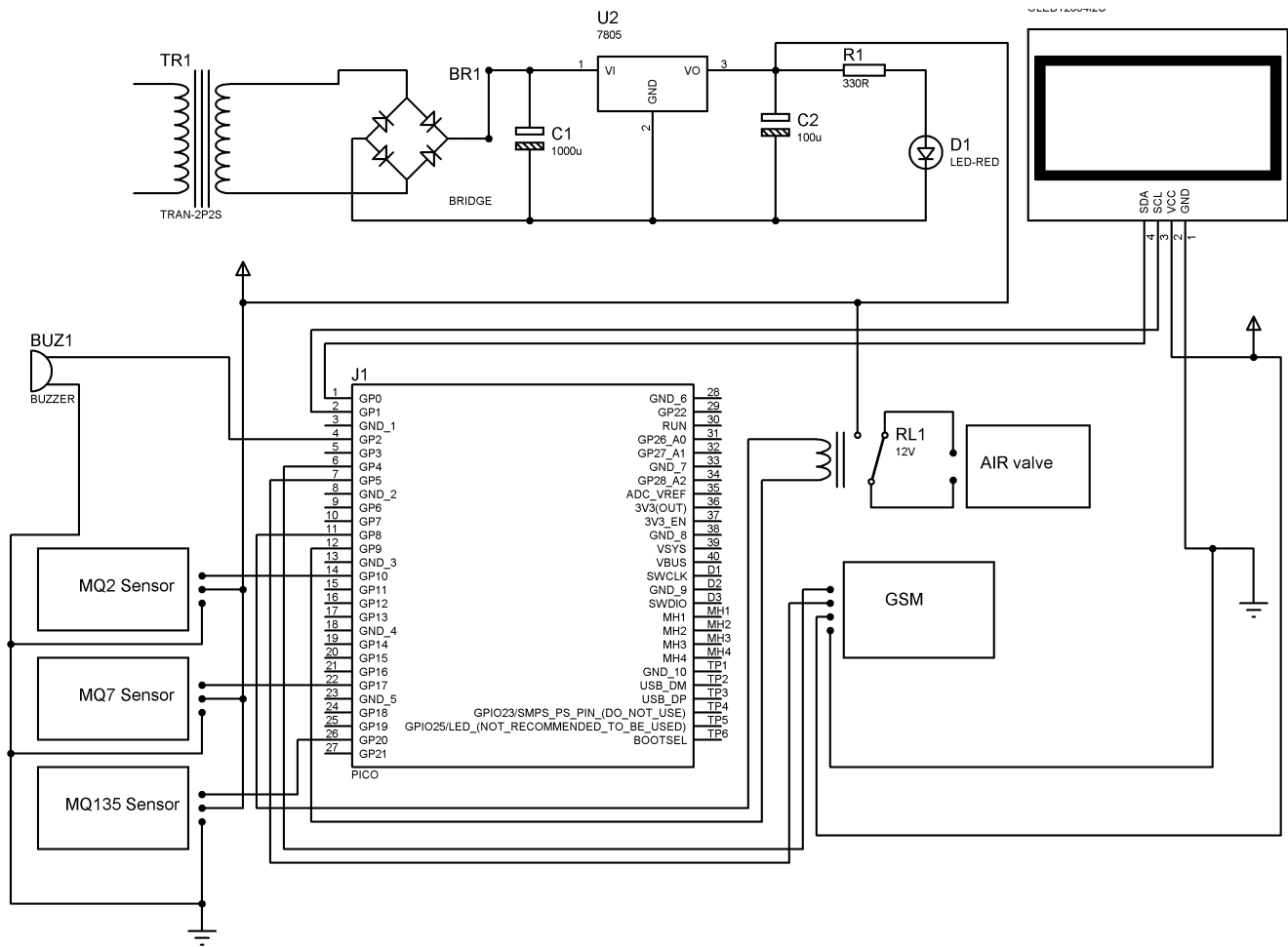


Fig 4.20: PM2.5 Sensor Module

V. Schematic Diagram



VI. Result

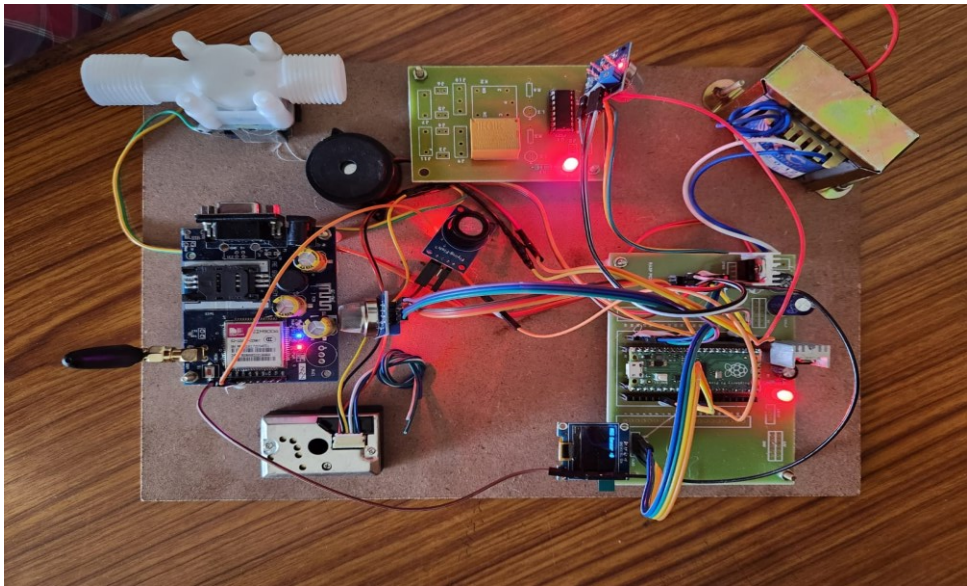


Fig- : Hardware Kit When in ON condition



Fig OLED displaying normal state of gas sensors

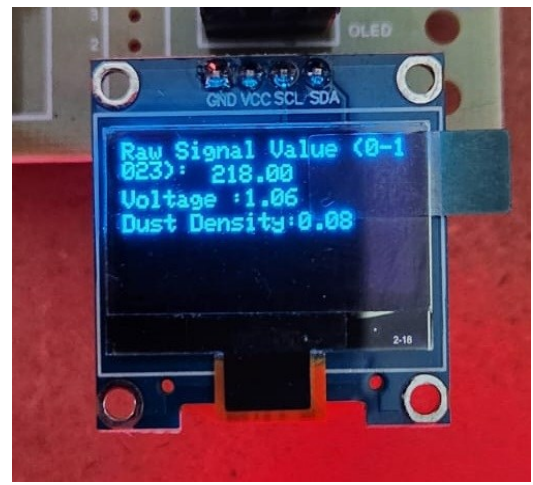


Fig: display of dust density



Fig OLED displaying leak status

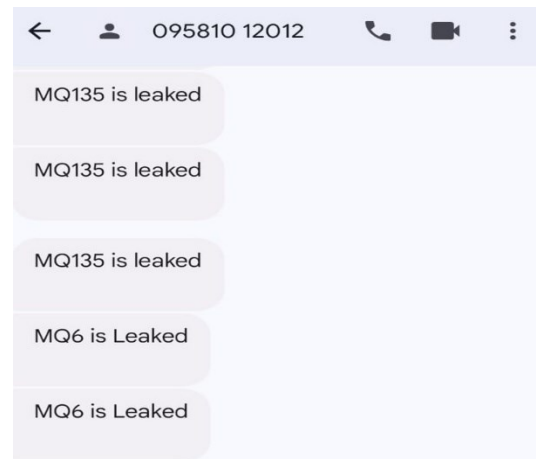


Fig: Status of gas detection through SMS

Results:

The successful development of the Digital Twin system for PM2.5 estimation and action policies resulted in the following:

1. High-Resolution PM2.5 Monitoring:

The system accurately detected PM2.5 and other air pollutants using sensors (MQ2, MQ7, MQ135, and

PM2.5 Dust Particle Sensor).

Enabled real-time monitoring of air quality with high precision.

2. Sensory Data Fusion:

Integrated multiple sensors for comprehensive data collection on pollutants like CO, NO₂, and PM_{2.5}.

Used sensory data fusion to recommend actionable policies based on pollutant levels.

3. Efficient Data Processing:

The Raspberry Pi Pico microcontroller processed sensory inputs in real time to estimate pollution trends.

Software (Embedded C, Micro Python, Arduino IDE) ensured seamless data handling and system control.

4. Stable Connectivity:

The Wi-Fi and GSM modules maintained robust communication with cloud platforms, enabling dynamic updates.

5. Interactive Feedback Mechanism:

The OLED display showcased pollutant levels and system status for immediate feedback.

The buzzer alerted users during critical pollution levels.

6. Adaptive Control Features:

Implemented an air valve motor to regulate airflow based on sensor inputs, improving air quality.

Recommended actionable policies dynamically, ensuring effective interventions.

7. Scalability and Expansion Potential:

The system design supports the addition of new sensors and modules to monitor more pollutants or expand functionalities.

8. Energy Efficiency:

Operated with low power consumption, leveraging an efficient power supply design.

9. Policy Decision Support:

Generated high-resolution air pollution maps and analytics to assist policymakers in creating evidence-based action plans.

Alignment of Sustainable Development Goals (SDG's)

SDG 3: Good Health and Well-being



1. Tracks PM2.5 levels to identify pollution hotspots and reduce health risks.
2. Uses data to predict health trends and improve preventive care.
3. Raises public awareness for safer outdoor activities

SDG 11: Sustainable Cities and Communities



1. Helps design eco-friendly urban spaces using air quality data.
2. Provides insights for traffic and industrial emission control.
3. Enhances city readiness for pollution-related emergencies.

SDG 13: Climate Action



1. Supports policy-making to cut pollution and carbon emissions.
2. Guides clean energy shifts based on air quality insights.
3. Promotes collaboration for integrated pollution control.

VIII. Future Scope:

1. Advanced Data Integration
 - Leverage cutting-edge satellite technologies for precise pollutant mapping.
 - Incorporate IoT-enabled sensors for real-time data collection and monitoring.
2. AI-Powered Analytics
 - Utilize deep learning models for improved and dynamic PM2.5 predictions.
 - Develop adaptive AI systems for enhanced environmental modeling.
3. Comprehensive Environmental Modeling
 - Expand the framework to include additional factors like noise and humidity.
 - Improve data fusion techniques to gain broader and deeper insights.
4. Localized Interventions
 - Design region-specific pollution management strategies for greater impact.
 - Refine policies and actions using local environmental data patterns.
5. Smart Cities Integration
 - Integrate DT systems with traffic management for better pollution control.
 - Optimize industrial activities based on real-time environmental data.

6. Global Collaboration

- Foster international partnerships to unify air pollution mitigation efforts.
- Align Digital Twin frameworks with global environmental standards.

7.Public Health Synergy

- Predict real-time health risks associated with air pollution levels.
- Guide preventive healthcare measures through integrated data insights.

8.Ecosystem Management

- Apply Digital Twin technology to monitor and improve water quality.
- Address soil pollution for broader ecosystem restoration and management.

S

References

- [1] Who Global Air Quality Guidelines: Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. Accessed: Feb. 15, 2022.
- [2] G. Hoek, R. M. Krishnan, R. Beelen, A. Peters, B. Ostro, B. Brunekreef, and J. D. Kaufman, "Long-term air pollution exposure and cardio—Respiratory mortality: A review," *Environ. Health*, vol. 12, no. 1, p. 43, Dec. 2013.
- [3] R. D. Brook, S. Rajagopalan, C. A. Pope, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, Y. Hong, R. V. Luepker, M. A. Mittleman, A. Peters, D. Siscovick, S. C. Smith, L. Whitsel, and J. D. Kaufman, "Particulate matter air pollution and cardiovascular disease," *Circulation*, vol. 121, no. 21, pp. 2331–2378, 2010.
- [4] Y.-L. Zhang and F. Cao, "Fine particulate matter (PM_{2.5}) in China at a city level," *Sci. Rep.*, vol. 5, no. 1, p. 14884, Oct. 2015.
- [5] Y. Chu, Y. Liu, X. Li, Z. Liu, H. Lu, Y. Lu, Z. Mao, X. Chen, N. Li, M. Ren, F. Liu, L. Tian, Z. Zhu, and H. Xiang, "A review on predicting ground PM_{2.5} concentration using satellite aerosol optical depth," *Atmosphere*, vol. 7, no. 10, p. 129, Oct. 2016. [Online]. Available: <https://www.mdpi.com/2073-4433/7/10/129>
- [6] M. Z. Joharestani, C. Cao, X. Ni, B. Bashir, and S. Talebiefandarani, "PM_{2.5} prediction based on random forest, XG Boost, and deep learning using multisource remote sensing data," *Atmosphere*, vol. 10, no. 7, p. 373, Jul. 2019.
- [7] J. Ma, Z. Yu, Y. Qu, J. Xu, and Y. Cao, "Application of the XG Boost machine learning method in PM_{2.5} prediction: A case study of Shanghai," *Aerosol Air Quality Res.*, vol. 20, no. 1, pp. 128–138, 2020.
- [8] A. Masood and K. Ahmad, "A model for particulate matter (PM_{2.5}) prediction for Delhi based on machine learning approaches," *Proc. Comput. Sci.*, vol. 167, pp. 2101–2110, Jan. 2020.
- [9] M. D. Yazdi, Z. Kuang, K. Dimakopoulou, B. Barratt, E. Suel, H. Amini, A. Lyapustin, K. Katsouyanni, and J. Schwartz, "Predicting fine particulate matter (PM_{2.5}) in the greater London area: An ensemble approach using machine learning methods," *Remote Sens.*, vol. 12, no. 6, p. 914, 2020.
- [10] T. Li, M. Hua, and X. Wu, "A hybrid CNN-LSTM model for forecasting particulate matter (PM_{2.5})," *IEEE Access*, vol. 8, pp. 26933–26940, 2020.
- [11] U. Pak, J. Ma, U. Ryu, K. Ryom, U. Juhyok, K. Pak, and C. Pak, "Deep learning-based PM_{2.5} prediction considering the spatiotemporal correlations: A case study of Beijing, China," *Sci. Total Environ.*, vol. 699, Jan. 2020, Art. no. 133561.
- [12] Y.-C. Chen and D.-C. Li, "Selection of key features for PM_{2.5} prediction using a wavelet model and RBF-LSTM," *Int. J. Speech Technol.*, vol. 51, no. 4, pp. 2534–2555, Apr. 2021.

- [13] H. Chang-Hoi, I. Park, H.-R. Oh, H.-J. Gim, S.-K. Hur, J. Kim, and D.-R. Choi, "Development of a PM2.5 prediction model using a recurrent neural network algorithm for the Seoul metropolitan area, republic of Korea," *Atmos. Environ.*, vol. 245, Jan. 2021, Art. no. 118021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1352231020307548>
- [14] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2017, pp. 4768–4777.
- [15] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why do tree-based models still outperform deep learning on tabular data?" 2022, arXiv:2207.08815.
- [16] M. Grieves, "Digital twin: Manufacturing excellence through virtual factory replication—A whitepaper by dr. Michael grieves," Florida Inst. Technol., Melbourne, FL, USA, White Paper 1–7, 2015.
- [17] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in *Transdisciplinary Perspectives on Complex Systems*. Berlin, Germany: Springer, 2017, pp. 85–113.
- [18] A. E. Saddik, "Digital twins: The convergence of multimedia technologies," *IEEE Multimed.*, vol. 25, no. 2, pp. 87–92, Apr./Jun. 2018.
- [19] E. Glaessgen and D. Stargel, "The digital twin paradigm for future NASA and US air force vehicles," in *Proc. 53rd AIAA/ASME/ASCE/AHS/ASC Struct., Struct. Dyn. Mater. Conf. 20th AIAA/ASME/AHS Adapt. Struct. Conf.*, 2012, p. 1818, doi: 10.2514/6.2012-1818.
- [20] Y. Liu, L. Zhang, Y. Yang, L. Zhou, L. Ren, F. Wang, R. Liu, Z. Pang, and M. J. Deen, "A novel cloud-based framework for the elderly healthcare services using digital twin," *IEEE Access*, vol. 7, pp. 49088–49101, 2019.
- [21] K. Reifsnider and P. Majumdar, "Multiphysics stimulated simulation digital twin methods for fleet management," in *Proc. 54th AIAA/ASME/ASCE/AHS/ASC Struct., Struct. Dyn., Mater. Conf.*, Apr. 2013, p. 1578.
- [22] C. Pylaniadis, S. Osinga, and I. N. Athanasiadis, "Introducing digital twins to agriculture," *Comput. Electron. Agricult.*, vol. 184, May 2021, Art. no. 105942.
- [23] B. Ketzler, V. Naserentin, F. Latino, C. Zangelidis, L. Thuvander, and A. Logg, "Digital twins for cities: A state-of-the-art review," *Built Environ.*, vol. 46, no. 4, pp. 547–573, Dec. 2020.
- [24] Y. Jiang, S. Yin, K. Li, H. Luo, and O. Kaynak, "Industrial applications of digital twins," *Philos. Trans. Roy. Soc. A*, vol. 379, no. 2207, 2021, Art. no. 20200360.
- [25] R. Rosen, G. Von Wichert, G. Lo, and K. D. Bettenhausen, "About the importance of autonomy and digital twins for the future of manufacturing," *IFAC-Papers OnLine*, vol. 48, no. 3, pp. 567–572, 2015.
- [26] A. K. Aap, M. S. Aap, R. Sabha, and L. Sabha, "Anil Baijal, IAS [4]," *Population*, p. 6, 2011.
- [27] S. K. Sahu and S. H. Kota, "Significance of PM2.5 air quality at the Indian capital," *Aerosol Air Quality Res.*, vol. 17, no. 2, pp. 588–597, 2017.
- [28] Q. Koziol, HDF5. Boston, MA, USA: Springer, 2011, pp. 827–833, doi: 10.1007/978-0-387-097664_44.
- [29] India Expo Centre & Mart Greater Noida: Information About Delhi. UNCCD. Accessed: May 8, 2022. [Online]. Available: <http://unccdcop14india.gov.in/about-delhi>
- [30] S. Moritz and T. Bartz-Beielstein, "ImputeTS: Time series missing value imputation in R," *R J.*, vol. 9, no. 1, pp. 207–218, 2017.
- [31] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi:10.1214/aos/1013203451.
- [32] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 2623–2631, doi: 10.1145/3292500.3330701.
- [33] T. Al Hosari, A. Al Mandous, Y. Wehbe, A. Shalaby, N. Al Shamsi, H. Al Naqbi, O. Al Yazeedi, A. Al Mazroui, and S. Farrah, "The UAE cloud seeding program: A statistical and physical evaluation," *Atmosphere*, vol. 12, no. 8, p. 1013, Aug. 2021. [Online]. Available: <https://www.mdpi.com/2073-4433/12/8/1013>
- [34] S. Ma, M. Zhang, J. Nie, J. Tan, B. Yang, and S. Song, "Design of double-component metal–organic framework air filters with PM2.5 capture, gas adsorption and antibacterial capacities," *Carbohydrate Polym.*, vol. 203, pp. 415–422, Jan. 2019.
- [35] Z. Wei, Q. Su, X. Wang, S. Long, G. Zhang, Q. Lin, and J. Yang, "Nanofiber air filters with high-temperature stability and superior chemical resistance for the high-efficiency PM2.5 removal," *Ind. Eng. Chem. Res.*, vol. 60, no. 27, pp. 9971–9982, Jul. 2021, doi: 10.1021/acs.iecr.1c01821.

- [36] A. Tobías, C. Carnerero, C. Reche, J. Massagué, M. Via, M. C. Minguillón, A. Alastuey, and X. Querol, “Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic,” *Sci. Total Environ.*, vol. 726, Jul. 2020, Art. no. 138540.