Indoor Air Quality and Utility Metering Platform for Predictive Building Decarbonization

Jan Tomlain[1], Juraj Tomlain[2]

[1] Dep. of Measurement, Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia

[2] T-Industry, s.r.o., Slovakia

1st tomlajan@fel.cvut.cz; 2nd duro@tind.sk

[1] https://orcid.org/0000-0002-3743-8394

[2] https://orcid.org/0009-0005-0638-8696

*Corresponding Author: tomlajan@fel.cvut.cz

HOW TO CITE:
Jan Tomlain, Juraj Tomlain (2026). Indoor Air Quality and Utility Metering Platform for Predictive Building Decarbonization. International Journal of Special Education, 41(12s), 395-403

ABSTRACT:
This paper presents a partial outcome of the REBECCA decarbonization project at the Slovak University of Technology in Bratislava, proposing a data platform where indoor air-quality (IAQ) sensing becomes a primary feature set for predictive analytics in building decarbonization. Moving beyond a hardware description, we detail an end-to-end pipeline from PoE-based sensing to AI-ready time-series data. Modular nodes measure CO2, particulate matter (PM), temperature and relative humidity, with ambient light used to derive occupancy proxies and segment operating regimes. IAQ dynamics are fused with electricity and gas meter streams to form cross-domain labels for ventilation demand and HVAC load. Data is transported via MQTT over TCP/IP to an in-house Remote Terminal Unit (RTU) that performs secure aggregation, buffering, automated quality checks, and time alignment before storing unified signals in a time-series database. This structure enables downstream models such as anomaly detection for ventilation faults, short-horizon forecasting for demand-flexibility windows, and interpretable drivers analysis that links IAQ excursions to schedule patterns and utility peaks. Initial operation across an academic semester showed stable connectivity and measurement consistency and surfaced repeatable IAQ events suitable for training and benchmarking models. The contribution is a reusable architecture and methodological baseline for integrating IoT, data analytics, and optimization, providing the measurement backbone for the next REBECCA phases targeting carbon-aware control strategies that reduce energy use while preserving healthy indoor environments.

Keywords: building energy monitoring, building decarbonization, IoT data platform.

I. Introduction

University buildings represent a challenging context for decarbonization. Occupancy is typically schedule-driven, space usage is heterogeneous—ranging from teaching rooms to offices and laboratories—and comfort and health requirements often conflict with energy-reduction objectives [1], [2]. In addition, the operational profile of university buildings varies significantly over time, reflecting the academic calendar. Teaching semesters, examination periods and summer breaks are associated with markedly different occupancy levels, ventilation demands and thermal loads [3]. As a result, building operation cannot be characterized by a single steady-state regime. In practice, many energy-optimization initiatives fail because the underlying data infrastructure is insufficient to capture this temporal and functional variability [4]. Without continuous, reliable and fine-grained measurements across different operational phases of the academic year, it is difficult to identify the sources of inefficiency, to assess the impact of implemented measures, or to train predictive models that generalize across seasons, operating regimes and space types [3], [4].

Indoor air-quality (IAQ) sensing provides essential operational context that complements conventional energy metering. In addition to commonly monitored parameters such as CO2 concentration and particulate matter (PM), the platform also measures indoor air temperature and relative humidity, which are critical indicators of thermal comfort, ventilation performance and building operation [5]. CO2 dynamics offer indirect evidence of occupancy levels and ventilation adequacy [6], [7], while PM measurements can reveal event-driven exposure patterns and the effectiveness of air exchange [8], [9]. Temperature and humidity measurements, in turn, provide insight into sensible and latent thermal loads and into the interaction between ventilation, heating and cooling systems under different seasonal and occupancy conditions [5], [10]. Together, these parameters form a coherent description of the indoor environment that cannot be inferred from energy consumption data alone [1].

When IAQ signals—including carbon dioxide, PM, temperature and relative humidity—are aligned with electricity and gas metering data, and with on-site photovoltaic (PV) generation where available, they create a powerful explanatory layer for understanding building energy behaviour [10], [11]. This integrated view enables a clearer distinction between occupancy-driven ventilation demand, thermally driven energy use and baseline consumption. Importantly, it also allows differences between academic periods—such as full teaching operation, reduced activity during examination periods, and low-occupancy summer operation—to be explicitly captured and analysed [3]. Such differentiation supports data-driven discussions on ventilation scheduling, heating strategies, thermal comfort management and future heat-recovery or recuperation concepts, particularly in buildings with mixed teaching, office and laboratory functions [11], [12].

This paper reports an initial deployment of such a monitoring and data-integration platform at the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, located on Technicka 5 in Bratislava and associated with the High Voltage Laboratory (LVN) and renewable-energy research infrastructure. The LVN site includes extensive experimental capabilities, an installed photovoltaic plant with a peak capacity of 20~kW, and air-to-water heat-pump equipment [1]. These characteristics make the facility a relevant real-world testbed for combined energy and environmental monitoring and align directly with the objectives of the REBECCA project, which aims to develop and experimentally verify data-driven methodologies for the decarbonization of university buildings at low technology readiness levels (TRL~1--3). Within this context, the presented platform provides the measurement backbone required to capture seasonal and academic-cycle variability and to support subsequent modelling, simulation and validation activities envisaged in the project.

The contributions of this work are fourfold. First, it provides a deployment description of a multi-room IAQ and energy-metering platform in a university environment, based on thirteen IAQ units with configurable sampling intervals between one and fifteen minutes and continuous operation across the academic year. Second, it presents an end-to-end data-platform design that enables secure ingestion, buffering, quality flagging and time alignment across heterogeneous data streams, in line with the data-engineering requirements of the REBECCA project [4]. Third, it introduces a cross-domain labelling strategy that links IAQ excursions and occupancy proxies to electricity and gas load regimes and to PV self-consumption windows under different operational phases of building use [3]. Finally, it defines a model-readiness baseline, outlining the technical prerequisites that must be met before predictive AI methods for building decarbonization can be meaningfully applied in real campus conditions [12], [13].

Fig. 1 Indoor air quality unit

II. Site Context and Monitoring Scope

A. FACILITY AND OPERATIONAL CONTEXT

The deployment is situated in an STU FEI facility on Technicka 5, hosting the High Voltage Laboratory (LVN) and related research laboratories, including renewable energy research infrastructure. Public descriptions of the LVN indicate a combination of high voltage testing facilities and renewable energy systems, such as photovoltaic installation with a peak capacity of 20 kW and an air to water heat pump rated for both heating and cooling. This combination illustrates the site's relevance for integrated energy monitoring and decarbonization experimentation.

A. Energy systems and measured energy flows

From an energy systems perspective, the platform integrates several key flows. Electricity consumption is captured through automated readings from electricity meters at the building level and, where available, from selected sub meters. Gas consumption is monitored through automated readings from gas meters associated with the on-site gas driven water heating appliances, which serves as the primary heating source. During the main heating season, space heating is provided predominantly by the gas boiler, while during shoulder periods the air to water heat pump is used to support heating and to enable experimentation with lower carbon operating modes. In addition, a photovoltaic system with a peak capacity of 20 kW is installed on the site and operated in a self-consumption, no export regime, allowing locally produced electricity to be utilized by other facilities in campus.

B. IAQ monitoring scope

Thirteen indoor IAQ units (device shown in Fig. 1) are installed across representative spaces, including classrooms Fig. 3, offices, and laboratories Fig. 4.

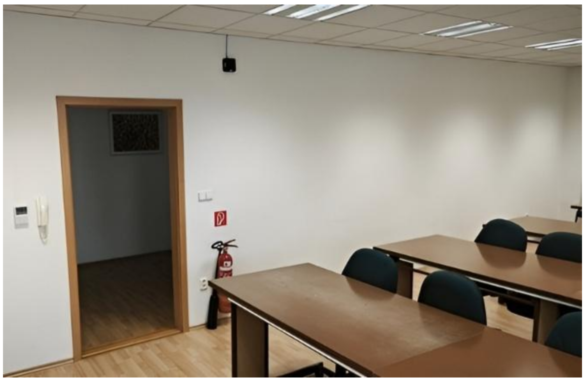

Fig. 2 IAQ unit in classroom

Fig. 3 IAQ unit in laboratory

Classrooms are characterized by schedule driven occupancy and pronounced ventilation demand peaks during teaching hours. Offices typically exhibit longer periods of stable occupancy and different ventilation patterns, while laboratories show higher variability due to activity driven PM events and stricter ventilation constraints. Each IAQ unit measures CO2 humidity, with a configurable sampling interval ranging from one minute to fifteen minutes depending on space type and monitoring.

III. System Architecture (Data Platform)

The complete diagram of the system architecture is shown in Fig. 2.

A. Sensing layer (IAQ units)

The IAQ units act as edge endpoints that generate continuous time series data at the room level [6], [8]. In addition to providing descriptive information about indoor environmental conditions, these measurements are explicitly treated as constraints and drivers for future energy optimization, for example in ventilation control or heat recovery strategies [10], [11].

B. Connectivity and ingestion

Telemetry is delivered using a lightweight IoT communication stack. MQTT over TCP/IP is employed to support low-overhead and scalable publish--subscribe ingestion, enabling efficient transmission of time-series data from distributed sensing nodes [14], [15]. Physical connectivity depends on the available local infrastructure. While the IAQ devices support wireless connectivity via Wi-Fi, the primary and preferred deployment option is wired Ethernet with Power over Ethernet (PoE) supply.

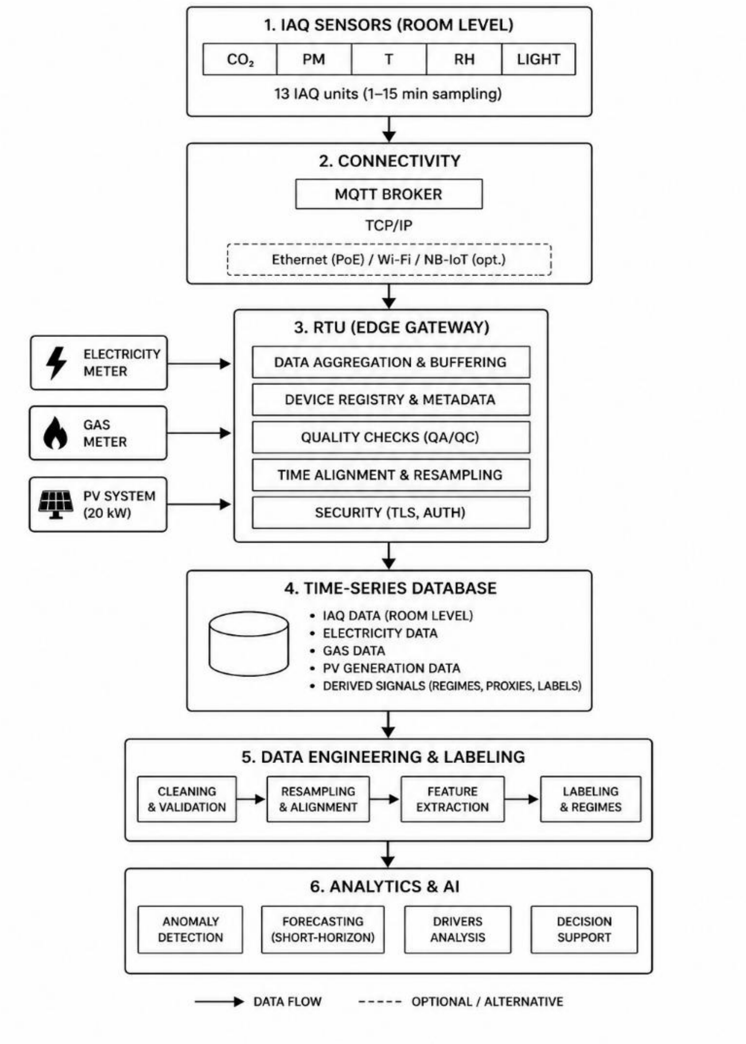

Fig. 2 System architecture

The use of PoE allows simultaneous data transmission and power delivery over a single cable, which reduces installation complexity and improves operational reliability in environments where Ethernet and PoE infrastructure is already present. Wireless connectivity is therefore used mainly as a complementary option in locations where cabling is impractical. In addition, the platform architecture is designed to accommodate NB-IoT as an optional extension for selected locations with limited or no access to fixed network infrastructure [16].

C. RTU orchestration and governance

A dedicated Remote Terminal Unit (RTU) provides the core platform functions required for AI readiness. These include data aggregation and buffering to tolerate outages and burst traffic, a device registry with metadata mapping devices to rooms and space types, and systematic quality checks covering plausibility ranges, missing data detection, and outlier flagging [4]. The RTU also performs time alignment by resampling heterogeneous streams to a canonical interval while preserving access to raw data, which is a prerequisite for reliable multi-source analytics [3].

D. Time series storage and access layer

A time series database stores unified data streams, including IAQ channels at the room level, electricity and gas meter series, PV generation series where available, and derived regime indicators such as occupancy proxies, heating regimes, and PV regimes [17], [18]. The database exposes consistent query interfaces that support both dashboarding and downstream analytics pipelines, enabling scalable analysis of high-resolution building and energy data [4].

IV. Data Engineering for Predictive Analytics (Ai Readiness)

To avoid "garbage in, garbage out"' in predictive modelling, the platform applies explicit data engineering practices [4]. Each sample is annotated with quality information rather than being destructively filtered, which is a commonly recommended approach for data-driven building analytics [3].

Plausibility checks are applied to CO2, PM, temperature, and humidity values, missing data are detected and quantified, and PM spikes are preserved as candidate physical events while potential sensor anomalies are flagged [8]. Long-term baseline shifts are tracked to inform later calibration and sensor maintenance [5].

Because the various streams operate at different sampling rates—IAQ between one and fifteen minutes and utility meters typically at coarser intervals—the platform defines canonical analytics buckets, such as fifteen-minute intervals, and consistent alignment rules [4]. Timestamp conventions are standardized using device timestamps with RTU receipt times as a fallback, which is essential for reliable multi-source time-series analysis [17].

On top of the aligned data, a practical and explainable labelling strategy is applied. IAQ excursion labels capture sustained CO2 exceedances above configurable thresholds and short-term PM spikes relative to baseline [6], [7]. Occupancy regimes are inferred using privacy-preserving proxies based on CO_2 slopes and periodicity, optionally supported by timetable information in later phases [5].

Energy-related regimes include electricity and gas peak windows identified relative to adaptive baselines, PV generation regimes distinguishing generating and non-generating periods, and heating regimes differentiating gas-boiler-dominant and heat-pump-dominant operation during shoulder seasons [10], [11]. These labels enable downstream anomaly detection, short-horizon forecasting, and interpretable drivers analysis linking occupancy, energy use, and IAQ outcomes [12], [13].

Deployment, Current Status on Site

The current deployment consists of thirteen indoor air-quality (IAQ) units installed across representative classrooms, offices and laboratory spaces within the monitored building. The units operate continuously with configurable sampling intervals between one and fifteen minutes, allowing the monitoring configuration to be adapted to different room functions and activity levels while maintaining consistent long-term data coverage. The data visualization from the IAQ units – CO2 levels are shown in ne plot in Fig. 5.

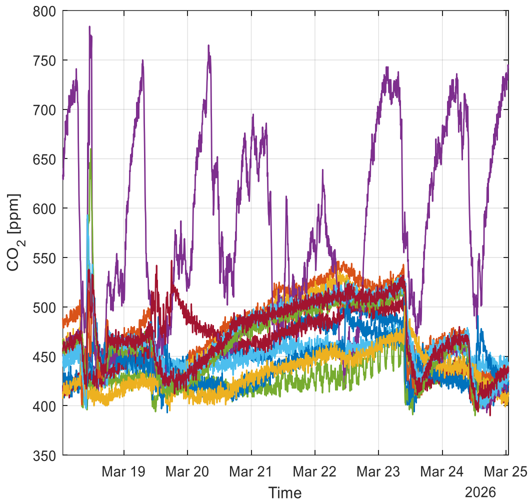

Fig. 4 CO_2 concentration over semester week

Electricity and gas meters are integrated into the same data platform, providing a unified view of electrical and thermal energy consumption. In addition, the building is equipped with an on-site photovoltaic system with a peak installed capacity of 20 kW, which is operated in a self-consumption (no-export) regime. PV production is therefore considered an integral contextual signal when analysing electricity demand and operational patterns. The PV production starts to be more relevant during summer and mid-summer part of the year, while the spring month like March is producing around 70 kWh of energy during overcast days as shown in Fig. 7.

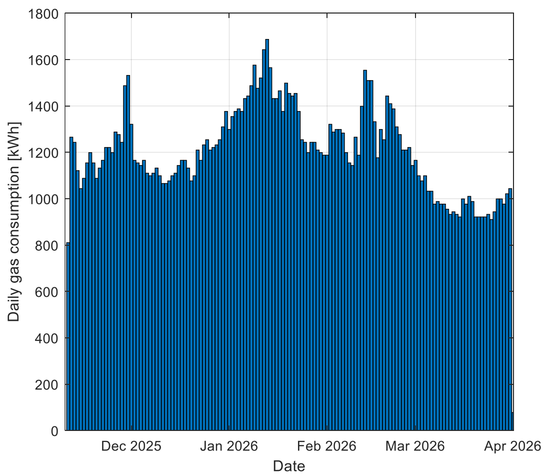

Fig. 5 Daily gas consumption during heating period

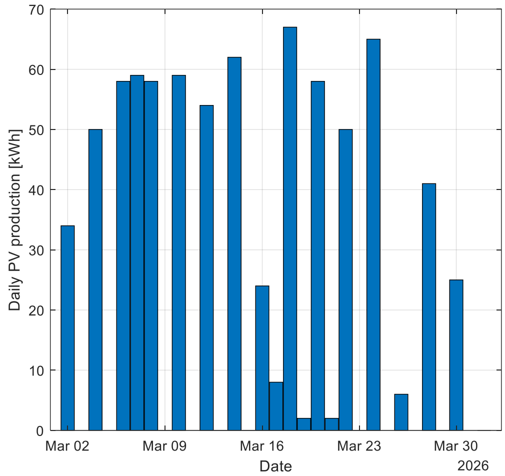

Fig. 6 Solar electricity production during spring month

Space heating is provided primarily by a gas boiler plant, which represents the dominant heat source during the main heating season. The gas consumption over the heating part of the year is shown in Fig. 6. During shoulder periods, an air-to-water heat pump supplements the heating supply, introducing an additional operating regime with different efficiency characteristics. The coexistence of multiple space types, energy carriers and heating technologies motivates the integrated monitoring approach adopted in the platform and provides a realistic basis for subsequent analysis and evaluation of decarbonization measures.

V. Decarbonization Connection

The central motivation of this work is that predictive AI for building decarbonization is not feasible without high-quality, well-structured data [3], [4]. In campus environments, models must learn recurring patterns related to academic timetables, seasonal heating operation, on-site photovoltaic (PV) production and differences between space types such as classrooms, offices and laboratories [1], [2]. Capturing these patterns requires continuous multi-stream measurements, consistent metadata, reliable time alignment and explicit quality flags and labels [4]. The platform described in this paper represents a minimum viable data foundation that must be in place before any meaningful modelling or optimization effort can begin [3].

Table 1: Indicative Intensity of Selected Energy Carriers

Energy carrier	CO2 / kWh equivalent intensity
Natural gas (combustion for heat)	~ 0.20 kg
Electricity (average European mix)	~ 0.24 – 0.30 kg
Solar PV (lifecycle average)	~ 0.03 kg

To illustrate the relevance of accurate energy and environmental data, Table 1 provides indicative carbon-intensity values for the dominant energy carriers used in the monitored building. Although exact values depend on operating conditions and accounting methodology, these reference figures highlight the scale of emissions associated with different energy sources and the importance of distinguishing between them in data-driven analyses [19], [20]. These values indicate that both thermal energy from gas and electricity from the European grid carry a comparable carbon burden on a per-kilowatt-hour basis, with electricity offering a clear decarbonization advantage as the generation mix continues to shift toward renewables [20]. From a methodological perspective, this underlines why precise measurement and separation of gas consumption, electricity demand and on-site PV generation are essential [4]. Without such resolution, it is not possible to evaluate the real carbon impact of electrification strategies, PV self-consumption or changes in heating operation [1], [2].

By treating indoor air quality (IAQ) parameters such as carbon dioxide concentration, particulate matter, temperature and relative humidity as constraint signals alongside energy data, the platform mitigates the risk that energy-optimization strategies degrade indoor environmental quality [6], [7], [8]. This is particularly important in educational buildings, where comfort and health requirements must be maintained across varying occupancy regimes throughout the academic yearn [5]. In the longer term, the integrated dataset supports the evaluation of building modernization pathways such as intelligent heat recovery and recuperation, which aim to maintain adequate ventilation while avoiding unnecessary increases in heating demand [10], [11].

Within the context of the REBECCA project, the presented platform directly supports the project's objectives at low technology-readiness levels (TRL 1-3). It provides the empirical basis for developing and validating data-driven methodologies for decarbonization, for comparing alternative operational regimes, and for assessing the potential impact of future low-carbon technologies under realistic conditions [1], [2]. Importantly, the focus is not on immediate optimization, but on establishing a replicable measurement and data-engineering framework that can be transferred to other university buildings and, more broadly, to similar institutional facilities in the Slovak Republic and beyond [3].

Current limitations of the platform include the proxy-based nature of occupancy estimation, limited access to detailed HVAC control states, and the need for ongoing sensor calibration and drift management [5]. These limitations are acknowledged as part of an iterative development process and are consistent with the exploratory and methodological character of the REBECCA project. Addressing them progressively will further strengthen the platform's ability to support predictive modelling and informed decision-making for building decarbonization [12], [13].

Conclusion

This paper presented an AI ready monitoring platform deployed in an STU FEI facility on Technicka 5, integrating electricity and gas metering with thirteen IAQ units across classrooms, offices, and laboratories, and designed to incorporate PV self-consumption and heating regime information. The main contribution is not novel hardware, but a practical data platform and labeling methodology required to develop reliable predictive AI for building decarbonization while safeguarding indoor air quality.

Next steps include expanding spatial coverage and harmonizing sampling policies, improving calibration and metadata governance, integrating PV inverter telemetry and additional HVAC or recuperation signals as the building modernizes, and implementing and validating predictive models for anomaly detection, forecasting, and drivers analysis.

Beyond the technical and technological scope described in this work, building decarbonization cannot be achieved solely through monitoring, analytics, and automation. An equally important and inseparable dimension is the behavioral and educational aspect, addressing the carbon footprint generated by occupants through their everyday activities, work patterns, and operational decisions. Energy-aware behavior of staff and users—such as ventilation practices, equipment usage, scheduling of high-load activities, or responses to indoor air quality feedback—directly influences both emissions and indoor environmental quality. The presented monitoring platform therefore also establishes a foundation for data-driven awareness, engagement, and education, enabling transparent feedback on how human behavior interacts with building systems. Embedding such insights into training, organizational culture, and decision-making processes represents a critical next step toward sustainable operation, where technological optimization and informed human behavior jointly contribute to long-term decarbonization goals.

Acknowledgement

This work was funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I04-03-V02-00033.

References

M. Santamouris and K. Vasilakopoulou, "Present and future energy consumption of buildings: Challenges and opportunities towards decarbonisation," *Energy Build.*, vol. 236, p. 110784, 2021.

A. Levesque, R. C. Pietzcker, L. Baumstark, and G. Luderer, "Deep decarbonisation of buildings energy services through demand and supply transformations," *Environ. Res. Lett.*, vol. 16, no. 5, p. 054071, 2021.

T. Hong, S. C. Taylor-Lange, S. D'Oca, D. Yan, and S. P. Corgnati, "Advances in research and applications of energy-related occupant behavior in buildings," *Energy Build.*, vol. 116, pp. 694–702, 2016.

T. Ahmad and H. Chen, "Data-driven methods for building energy prediction and analysis," *Renew. Sustain. Energy Rev.*, vol. 82, pp. 297–316, 2018.

L. M. Candanedo and V. Feldheim, "Accurate occupancy detection of an office room from light, temperature, humidity and CO_2 measurements," *Energy Build.*, vol. 112, pp. 28–39, 2016.

A. Persily and L. de Jonge, "Carbon dioxide generation rates for building occupants," *Indoor Air*, vol. 27, no. 5, pp. 868–879, 2017.

O. A. Seppänen, W. J. Fisk, and M. J. Mendell, "Association of ventilation rates and CO_2 concentrations with health and performance," *Indoor Air*, vol. 9, no. 4, pp. 226–252, 1999.

L. Morawska and others, "Indoor air quality and the global burden of disease," *Environ. Int.*, vol. 118, pp. 159–168, 2018.

B. Zhao and J. Wu, "Particle deposition in indoor environments," *Atmos. Environ.*, vol. 41, no. 20, pp. 3862–3876, 2007.

S. Wang and Z. Ma, "Supervisory and optimal control of building HVAC systems," *Energy Build.*, vol. 40, no. 6, pp. 1063–1075, 2008.

F. Oldewurtel, A. Parisio, C. N. Jones, and others, "Use of model predictive control and weather forecasts for energy efficient building climate control," *Energy Build.*, vol. 45, pp. 15–27, 2012.

. Wei, Y. Wang, and Q. Zhu, "Deep reinforcement learning for building HVAC control," in *Proceedings of the ACM International Conference on Energy-Efficient Computing and Networking (e-Energy)*, 2017, pp. 1–10.

F. Ruelens, B. J. Claessens, S. Vandael, and others, "Residential demand response using reinforcement learning," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3131–3140, 2017.

R. A. Light, "Mosquitto: server and client implementation of the MQTT protocol," *J. Open Source Softw.*, vol. 2, no. 13, p. 265, 2017.

D. Thangavel, X. Ma, A. Valera, H. X. Tan, and C. K. Y. Tan, "Performance evaluation of MQTT and CoAP via a common middleware," *IEEE Ninth Int. Conf. Intell. Sens. Sens. Netw. Inf. Process.*, pp. 1–6, 2014.

U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 2, pp. 855–873, 2017.

A. Bader, O. Kopp, and M. Falkenthal, "Survey and comparison of open source time series databases," *Proc. 2017 Int. Conf. Big Data*, pp. 1–8, 2017.

S. Li and J. Wen, "Review of building energy modeling for control and operation," *Renew. Sustain. Energy Rev.*, vol. 105, pp. 174–194, 2019.

IPCC, *Climate Change 2014: Mitigation of Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press, 2014.

International Energy Agency, "CO_2 Emissions from Fuel Combustion 2022," IEA, Paris, France, 2022