Kalpa Publications in Computing
Volume 22, 2025, Pages 792–803
Proceedings of The Sixth International Conference on Civil and Building Engineering Informatics

A Self-Attention Fusion-Based BI-LSTM Framework for Occupant-Centric Prediction of Indoor Environmental Quality

Min Jae Lee[1] and Ruichuan Zhang[1]
[1] Virginia Polytechnic Institute and State University, Blacksburg, VA, US.
minjae0624@vt.edu, ruichuanz@vt.edu

Abstract

The integration of occupant data into the management of indoor environment factors is gaining increasing attention for creating intelligent and inclusive built environments. Existing approaches have mostly relied on static models, often failing to account for the ever-changing nature of occupant behavior and environmental factors across time and dimensions. Recent advancements in deep learning, especially deep sequential models capable of capturing both local and global dependencies between time steps, provide an opportunity to overcome these challenges. To address these challenges, the authors propose an LSTM-based model framework that utilizes multimodal self-attention-based fusion, real-time occupant data, indoor environmental quality (IEQ) data, and outdoor environmental data to predict future IEQ conditions, preferred IEQ conditions, and classify current IEQ conditions based on collected occupant feedback. To develop and test the proposed framework, four key steps were followed: (1) collecting IEQ data through smart sensors, (2) collecting perceived occupant feedback, (3) collecting outdoor environmental data, and (4) developing an attention-fusion-based Bi-Directional LSTM(Bi-LSTM) model. The proposed framework was tested at the Virginia Tech Blacksburg campus, showing promising results.

1 Introduction

Indoor Environmental Quality (IEQ) includes various aspects, including indoor air quality, lighting, thermal conditions, and acoustics, which are crucial in how occupants feel in indoor spaces (USGBC 2014). Furthermore, in today's modern society, humans have largely become an indoor species, spending up to 90% of their time indoors (EPA 2023). Many studies have found that well-maintained IEQ conditions result in improved occupant well-being, while sub-optimal IEQ conditions result in sick building syndromes and other health risks (Samet and Spengler 2003, Shan et al. 2018, Cincinelli and Martellini 2017, Jones 1999). Furthermore, buildings are becoming more dynamic and complex, especially with climate changes, which calls for integrating holistic and multiple factors into IEQ modeling and assessment (Smarra et al. 2018, Du et al. 2021, Liu et al. 2022).

Previous research in IEQ modeling and assessment for indoor environment management, while comprehensive, has been limited. Tien et al. (2022) stated that existing efforts for IEQ modeling and assessment often rely on traditional statistical models, which do not leverage the capabilities of deep learning to analyze complex interactions among multiple IEQ factors. Furthermore, Aldakheel (2023) reviewed the application of Artificial Intelligence (AI) in building management systems, noting that while there is a growing interest in using deep learning for energy efficiency and comfort, many studies have yet to explore its full potential in real-time IEQ monitoring and management. Second, previous studies have approached IEQ modeling and assessment using isolated data streams, neglecting the holistic integration of indoor and outdoor environmental factors. Lolli et al. (2022) stated that many existing models overlook the integration of indoor and outdoor data, which can result in an incomplete understanding of the indoor environment. Similarly, Appau (2023) conducted a post-occupancy evaluation of university student housing, highlighting that many studies focus solely on indoor conditions without considering the influence of external environmental factors, such as outdoor air quality and weather conditions. Furthermore, Sun et al. (2019) stated that the development of indoor air quality monitoring systems that often rely on isolated data streams can compromise the quality and reliability of the data collected.

To address these challenges, this paper proposed a novel framework that utilizes attention-based multimodal fusion with Bi-LSTM model architecture that can predict future IEQ conditions and classify current IEQ conditions from multidimensional IEQ data, outdoor environmental data, and occupant feedback. The proposed framework for IEQ prediction and classification consists of four main components: (1) collecting IEQ data through smart sensors, (2) collecting perceived occupant feedback, (3) collecting outdoor environmental data, and (4) developing an attention-fusion-based Bi-LSTM model to predict future IEQ conditions as well as classify IEQ conditions based on collected occupant feedback.

2 Background

2.1 Importance of IEQ

As humans spend more time in indoor spaces, reaching up to 90% of their time (EPA 2023), IEQ conditions are increasingly recognized as critical factors impacting occupant health and productivity within indoor environments. Recent literature highlighted the need for comprehensive and standardized IEQ data and measurement methods and emphasized that understanding of IEQ is still limited compared to outdoor air quality (Vardoulakis et al. 2020). Additionally, studies have shown that high concentrations of indoor air pollutants, such as CO_2, can lead to symptoms associated with Sick Building Syndrome (SBS), underscoring the importance of effective IEQ assessment in office environments (Shamsudin 2023). Zahaba et al. (2022) also showed that human exposure to indoor air pollutants can be significantly higher than that of outdoor pollutants, which raises concerns about the health impacts of poor IEQ. Furthermore, Deng et al. (2023) demonstrated that IEQ significantly affects occupants' productivity in indoor environments, particularly in the context of the COVID-19 pandemic, highlighting the need for improved IEQ management strategies (Deng et al., 2023). Furthermore, a study has also found a relationship between occupant mental health and the indoor environment, further highlighting the importance of well-maintained IEQ conditions (Beemer et al. 2019).

As the importance of well-maintained IEQ conditions increases, accurate modeling and assessment of IEQ conditions are also becoming important. Tagliabue et al. (2021) predicted future IEQ conditions following a data-driven approach with an IoT sensor network. Similarly, Lee and Zhang (2024) predicted future IEQ conditions at educational facilities utilizing multimodal occupant feedback along with multidimensional IEQ data. Kapoor et al. (2022) predicted future CO_2 concentration levels for office rooms using a machine learning-based approach. Fritz et al. (2022) evaluated machine learning models to classify occupants' perceptions of their indoor environment. IEQ models can be integrated into downstream applications such as smart building systems to enable occupant-centric control, where environmental settings are adjusted based on the predictions from IEQ models. Despite the advancements in predictive modeling and machine learning applications for IEQ, there remains a gap in fully understanding the complex relationship between IEQ and its effects on occupants along with outdoor environmental conditions

2.2 Deep Learning in Indoor Environments

The application of deep learning-based models in indoor environments has gained attraction, particularly for tasks such as indoor localization and environmental monitoring. Deep learning, which is a subset of machine learning, holds the improved ability to compute and classify large and complex datasets. Deep learning, when combined with multimodal data, can result in improved model performance (Meng et al. 2020). Recent studies have demonstrated the effectiveness of deep learning models in enhancing accuracy. For instance, Chenari et al. (2017) developed a CO_2- based demand-controlled ventilation strategy using deep learning algorithms, which not only improves IEQ but also optimizes energy consumption. Lee and Zhang (2024) developed an IEQ prediction model, based on Convolutional Neural Network (CNN), achieving high performance. Similarly, Rizk et al. (2019) showcased the use of deep learning models, specifically Convolutional Neural Networks (CNNs), for indoor localization based on Received Signal Strength (RSS) and Channel State Information (CSI), achieving notable improvements in localization accuracy. However, challenges remain regarding the generalization of these models across different indoor environments and the variability of data inputs, highlighting a need for further research to improve model robustness and applicability (Cretescu et al. 2019).

2.3 Multimodal Data Fusion

Multimodal data, collected from different sources and sensors, encompasses information that can show unique insights into each type of data as well as the overall system of interest (Lahat et al. 2016). Gao et al. (2020) surveyed deep learning for multimodal data fusion, emphasizing the importance of effectively combining different modalities to improve event understanding, especially when one modality is incomplete. Multimodal data fusion is increasingly recognized as a vital approach for improving IEQ assessments and modeling systems. Out of different algorithms for fusing multimodal data, attention-based fusion is gaining traction due to its ability to selectively focus on the most relevant features across modalities, thereby enhancing the overall model performance and resilience to missing data. Lin et al. (2023) fused multi-sensor data based on a self-attention mechanism. Similarly, Chan-To-Hing and Veeravalli (2024) proposed a cross-attention-based data fusion approach for Masked Autoencoders in remote sensing. Zhao et al. (2022) proposed an attention-based multimodal fusion model for human activity recognition. Liu et al. (2023) proposed a Transformer-based fusion model with modality-specific tokens to achieve effective cross- modal interaction. While multimodal data fusion is increasingly being applied across various domains, its application to IEQ or indoor environment modeling remains limited. Despite its potential to enhance the accuracy and comprehensiveness of indoor environment assessments, the use of attention-based fusion techniques in this field is still in its early stages, with only a few studies exploring these capabilities for effective IEQ prediction.

3 Proposed Framework

The proposed framework for IEQ prediction and classification consists of four main components (Figure 1): (1) collecting IEQ data through smart sensors, (2) collecting occupant feedback through a developed occupant feedback user interface (UI), (3) collecting outdoor environmental data through an API, and (4) developing an self-attention based Bi-LSTM model to predict future IEQ conditions as well as classify current IEQ conditions based on collected occupant feedback.

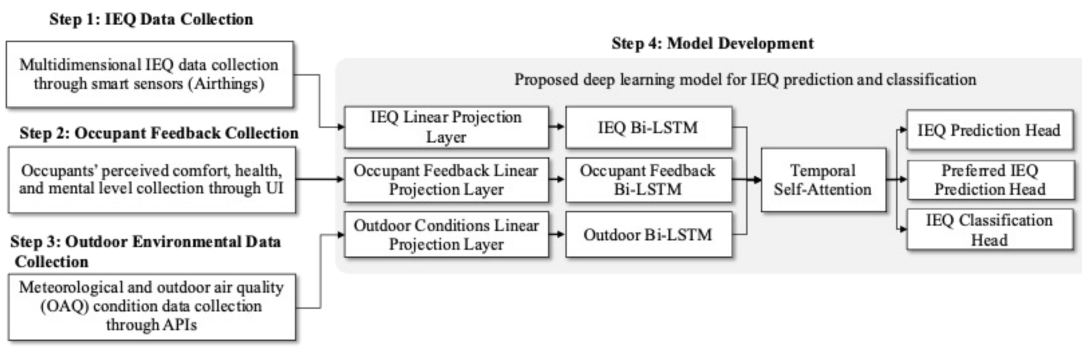

Figure 1: Proposed Framework

3.1 IEQ Data Collection

The first step of the proposed framework involves collecting real-time IEQ condition data using smart sensors, specifically the Airthings View Plus device. This sensor collects variable IEQ data, including temperature, humidity, CO_2 levels, VOC concentration levels, indoor air pressure, PM_1, $PM_{2.5}$, and Radon levels. All these variables are crucial for assessing and modeling indoor environmental conditions as well as occupant comfort. For example, high levels of CO_2 can indicate insufficient ventilation, leading to potential cognitive impairment, while VOCs and particulate matter (PM_1 and $PM_{2.5}$) are linked to respiratory issues and general discomfort. By capturing these diverse parameters continuously in real-time, the framework aims to provide a comprehensive dataset that reflects the dynamic nature of indoor environments.

3.2 Occupant Feedback Collection

The second step of the proposed framework includes collecting occupant feedback using an UI developed by the authors. This UI is designed to be accessible and intuitive, allowing occupants to indicate their perceived comfort, health, and mental level on a 5-point Likert scale. This subjective feedback serves as a valuable complement to the objective sensor data collected in the first step, as it helps bridge the gap between measured environmental conditions and their perceived impact on the occupants. By capturing individual perceptions, the framework can better account for variability in comfort preferences among different users, which may be influenced by factors such as age, health, or personal sensitivity to environmental changes. The collection of this feedback also allows for the identification of temporal patterns in occupant comfort, which can be analyzed alongside the objective IEQ data to create a more adaptive indoor environment that responds to the occupants' needs in real-time.

3.3 Outdoor Environmental Data Collection

The third step of the proposed framework involves collecting outdoor environmental data through an API, specifically the National Oceanic and Atmospheric Administration (NOAA) API. This provides additional context and perspective for the indoor environment. These data include outdoor temperature (°F), outdoor humidity (%), wind speed (mph), precipitation (in), cloud cover (%), $PM_{2.5}$ ($\mu g/m^3$), PM_{10} ($\mu g/m^3$), CO_2 (ppb), O_3 (ppb), and CO (ppm), all of which can influence IEQ and IAQ conditions as well as how occupants feel in indoor environments. Integrating outdoor condition data allows the model to consider external factors that can potentially impact the indoor environment, enabling predictions and classifications to be more comprehensive and holistic, and account for variations in outdoor environmental conditions.

3.4 Attention-Fusion based Bi-LSTM Model Development

The proposed framework's final step involves developing a Bi-LSTM with a temporal self-attention model for IEQ predictions and classifications. The proposed model is a deep learning-based architecture designed to handle temporal and multimodal data, incorporating Bi-LSTM layers and interpretable temporal attention mechanisms. This architecture allows the model to capture both local temporal patterns through the LSTM layers and global dependencies across time steps using the self-attention mechanism. The proposed model also includes three final heads: (1) predicting future IEQ conditions without considering occupant preferences, (2) predicting preferred IEQ conditions based on collected occupant feedback (i.e., comfort and health levels) using a weighted loss function, and (3) classifying IEQ conditions based on occupant feedback. This design allows the model architecture to simultaneously focus on current conditions, adjust for occupant preferences, and provide a classification of occupant well-being. The interpretable temporal self-attention mechanism enables the identification of influential features from the collected data that affect the predictions. The proposed model architecture consists of three key components:

1. **Data Processing Layers**: The model begins with linear projection layers that project each data modality (IEQ, occupant feedback, and outdoor conditions) into a common hidden dimension. This ensures consistency across all input data types and prepares each modality for further processing. The input sequences are then fed into separate Bi-LSTM layers for each modality, following the model fusion approach. The Bi-LSTM layers allow the model to capture temporal dependencies within each data stream, leveraging the bidirectional processing to incorporate both past and future data.

2. **Attention-Fusion Layers**: The outputs of the Bi-LSTM layers are then processed by temporal self-attention layers. These attention layers are applied separately to each modality, allowing the model to focus on important time steps within the data sequences. The self-attention mechanism helps identify which time steps in the sequence contribute the most to the prediction and classification tasks, capturing both short-term variations and long-term trends. The attention outputs from all three modalities (IEQ, outdoor, feedback) are then fused, combining the learned features into a unified representation. This fusion process involves stacking the attention outputs from each modality and averaging them, followed by a fully connected layer to reduce the dimensionality and integrate the information. Additionally, the model incorporates learnable weights for each modality, which allows the model to prioritize certain data streams based on their relevance to the task at hand.

3. **Prediction and Classification Heads**: The fused representation is passed through separate fully connected layers for both regression and classification tasks. The proposed model includes three output heads:
    a. **IEQ Prediction Head**: Predicts future IEQ conditions without considering occupant preferences, using a regression approach.
    b. **Preferred IEQ Prediction Head**: Predicts conditions that are more favorable to the occupants based on collected feedback, incorporating a weighted loss function to give higher importance to preferred conditions.
    c. **Classification Heads for Occupant Feedback**: Outputs predictions for occupant comfort and health levels, treating these as classification tasks.

4 Experiments and Results

4.1 Data Collection

To test the proposed framework, a set of experiments was conducted from an academic building on the Virginia Tech Blacksburg campus. The selected building represents a typical academic environment with varied occupant activities ranging from self-studying to listening to lectures. IEQ data were recorded every 2 minutes, resulting in a dataset capable of capturing subtle changes in environmental conditions. Occupant feedback was collected through an UI, and additional feedback data were augmented using an AI-based algorithm to increase the volume and diversity of the dataset for this experiment. Outdoor environmental data were collected from an API and augmented using a similar AI-based approach to provide more comprehensive coverage. After augmentation, the total number of data points for each data type was 23,636.

4.2 Data Processing

Several preprocessing steps were followed to ensure that the data was in the correct format and suitable for effective and accurate analysis as well as prediction and classification by the proposed model.

(1) Loading and Merging Datasets. The IEQ, outdoor, and occupant feedback datasets were loaded and merged based on a timestamp, double-checking that all modalities were aligned, resulting in a cohesive dataset for training the model. Furthermore, the recorded timestamp was converted to a consistent datetime format, and the merged dataset was sorted chronologically to maintain temporal dependencies.

(2) Data Normalization. All collected data, including IEQ/IAQ data, occupant feedback data, and meteorological and OAQ data, were normalized using the MinMaxScaler to bring all features to a common scale, allowing more effective learning by the proposed model.

(3) Imputation for Feedback. Missing feedback values were handled by first forward-filling to extend previously recorded feedback until new feedback was recorded. This ensured the continuity of feedback data, which is crucial for time series training. In cases where initial feedback was missing, a neutral value of 3 was used to fill the gaps.

(4) Sequence Generation and Data Splitting. Sequences of data were generated with a length of 15 timesteps to capture temporal patterns effectively. The target for regression (future IEQ values)

was set to 1 timestep ahead of the input sequence. The dataset was split into training, validation, and test sets in a 60-20-20 ratio to ensure proper evaluation of the model.

4.3 Evaluation Metrics

To evaluate the proposed model, five key metrics were used: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R^2, and accuracy. MSE, as shown in Equation 1, measures the average squared difference between the predicted and actual values, indicating how well the model predicts future IEQ conditions. A lower MSE suggests better performance. MAE, shown in Equation 2, measures the average magnitude of the errors without considering their direction, providing a straightforward interpretation of prediction accuracy. Compared to MSE, MAE is less sensitive to outliers, and lower MAE indicates better performance. RMSE, shown in Equation 3, is the square root of MSE and provides an interpretable error metric in the same units as the target variable. R^2, as shown in Equation 4, represents the proportion of variance in the dependent variable that is predictable from the independent variables, with a value closer to 1 indicating a better fit.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(T_i - \hat{T}_i)^2 \quad (1)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|T_i - \hat{T}_i\right| \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} \quad (3)$$

$$\mathrm{R}^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (4)$$

Accuracy, as shown in Equation 5, is used for classification tasks to assess the proportion of correct predictions made by the model out of all predictions.

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\text{Total number of samples}} \quad (5)$$

4.4 Ablation Analysis

The proposed model was tested in Python 3 using an NVIDIA A100 GPU. Comparative analyses were conducted to evaluate the impact of model architecture and data fusion strategies on the prediction and classification of IEQ. The analyses included two primary model configurations: (1) the proposed BiLSTM with Self-Attention, and (2) the standard LSTM with Self-Attention. These configurations were chosen to assess the ability to capture complex temporal dependencies and deliver accurate predictions. Additionally, the evaluation considered two different data fusion strategies: (1) Attention-Based Fusion, which utilized multi-head attention mechanisms to integrate IEQ, outdoor, and occupant feedback data, and (2) GAT-Based Fusion, which employed Graph Attention Networks to explicitly model relationships between data modalities.

The evaluation compared two key aspects: model architecture and fusion strategies. In terms of model architecture (Table 1), the proposed Bi-LSTM with Self-Attention demonstrated significantly better performance than the standard LSTM with Self-Attention. The Bi-LSTM achieved a 38.7% lower MSE, indicating its superior ability to capture complex temporal dependencies in the data, which is crucial for accurate indoor environmental quality (IEQ) predictions. Both models maintained perfect classification accuracy for comfort and health levels, but the proposed Bi-LSTM consistently provided more precise regression results, reflecting its improved predictive capabilities.

Model	MSE	% Diff. in MSE	MAE	RMSE	*R*2	Comfort Acc.	Health Acc.
Bi-LSTM with Self-Attention	**0.1515**	**–**	**0.2239**	**0.3889**	**0.8093**	**1.0000**	**1.0000**
LSTM with Self-Attention	0.2469	+38.7%	0.3569	0.4968	0.6424	1.0000	1.0000

Table 1: Performance difference between different model architectures

For fusion strategies (Table 2), Attention-based Fusion outperformed GAT-Based Fusion by 19.3% in terms of MSE. The use of multi-head attention allowed the model to effectively integrate IEQ, outdoor conditions, and occupant feedback data, resulting in more accurate predictions. While GAT Fusion was less effective than Attention- Based Fusion, it still offered meaningful benefits over other methods by modeling the relationships between data modalities explicitly. However, the superior performance of Attention-Based Fusion highlighted its capability to better capture and utilize the diverse information available across different data sources. Despite the differences in MSE, both fusion strategies showed strong classification results, with Attention-Based Fusion maintaining a slight advantage in terms of overall predictive accuracy.

Fusion Strategy	MSE	% Difference in MSE	MAE	RMSE	*R*2	Comfort Accuracy	Health Accuracy
Attention-Based Fusion	0.1515	–	0.2239	0.3889	0.8093	1.0000	1.0000
GAT-Based Fusion	0.1878	+19.3%	0.2908	0.4330	0.7483	0.2381	1.0000

Table 2: Fusion method comparison

The results showed that the proposed Bi-LSTM with Self-Attention and Attention-Based Fusion consistently delivered superior performance, achieving lower prediction errors and more accurate integration of data compared to the other models and fusion strategies. This confirms the effectiveness of the proposed approach in capturing temporal dependencies and utilizing diverse information for predicting and classifying IEQ conditions.

4.5 Predicted vs. Actual Values Plot

Figure 2 compares actual and predicted values for four IEQ parameters: Radon, CO_2, $PM_{2.5}$, and Humidity. The plots show that the predicted values closely follow the general trend of the actual data, indicating the model's ability to track fluctuations in the IEQ parameters. However, some deviations are observed, suggesting areas where prediction accuracy could be further refined. Overall, the model demonstrates its capacity to capture key patterns in the data across different IEQ variables.

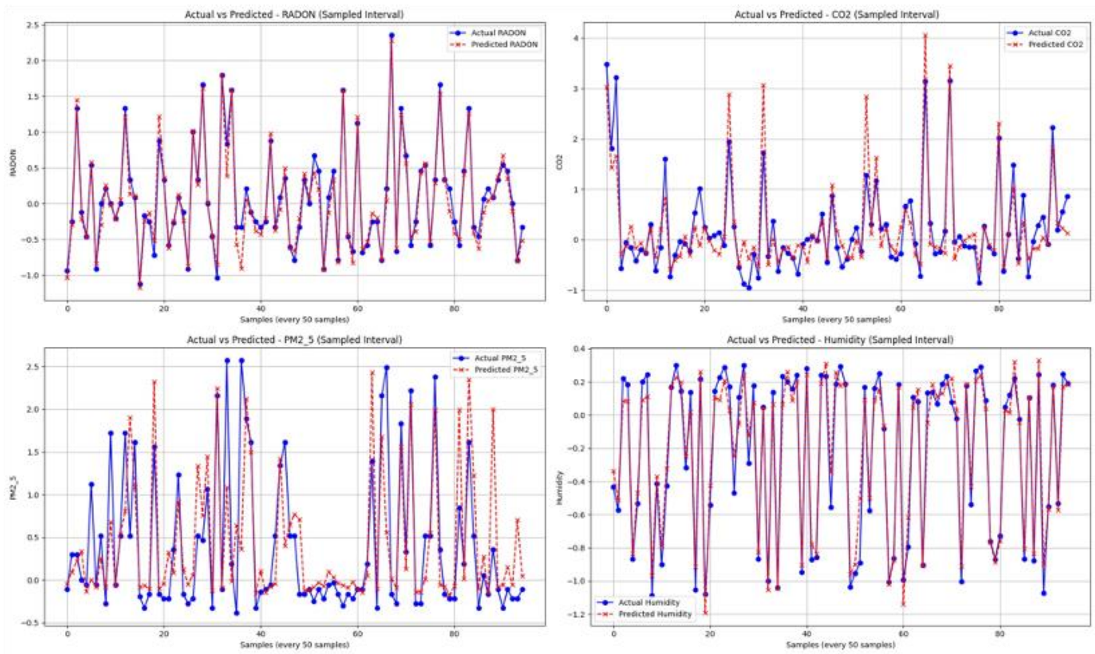

Figure 2: Predicted vs. Actual Values Plot

5 Conclusion

In conclusion, this paper presents a novel Bi-LSTM model with self-attention that integrates IEQ data, outdoor environmental data, and occupant feedback through attention-based fusion to predict and classify indoor environmental quality conditions. The proposed framework leverages the advantages of deep learning and multimodal data fusion to provide a comprehensive assessment of indoor environments. Results from the comparative analyses demonstrate that the proposed Bi-LSTM model with attention-based data fusion consistently outperforms the baseline models, highlighting the benefits of using advanced model architectures and fusion techniques. This research contributes to the growing body of work on IEQ assessment, offering valuable insights for improving occupant comfort, health, and overall indoor environment quality.

While the proposed framework's results are promising, several areas for future work remain. Additional data modalities, such as occupancy patterns, activity data, and human vital signs, could be incorporated to further enhance the model's capabilities. This will allow the model to capture the dynamic interactions between environmental conditions and occupant behavior more effectively, creating adaptive and personalized indoor environments. Future research could look into evaluating the scalability of the proposed framework across different building types, including residential, commercial, and mixed-use facilities, to assess its generalizability. Furthermore, integrating Model Predictive Control (MPC) could be explored to enable real-time, proactive adjustments of indoor environmental conditions based on the model's predictions. By employing MPC, the system would provide an intelligent, closed loop control mechanism to maintain optimal IEQ conditions continuously.

References

Aldakheel, J. (2023). *Indoor environmental quality evaluation of smart/artificial intelligence techniques in buildings – a review*. E3S Web of Conferences, 396, 01101. https://doi.org/10.1051/e3sconf/202339601101

Appau, W. (2023). *Indoor environmental quality and energy use intensity: an empirical post-occupancy evaluation test of on-campus university student housing in Ghana*. Property Management, 42(3), 333–352. https://doi.org/10.1108/pm-07-2023-0058

Beemer, C. J., Stearns-Yoder, K. A., Schuldt, S. J., Kinney, K. A., Lowry, C. A., Postolache, T. T., Brenner, L. A., & Hoisington, A. J. (2019). *A brief review on the mental health for select elements of the built environment*. Indoor and Built Environment. https://doi.org/10. 1177/1420326x19889653

Chan-To-Hing, H., & Veeravalli, B. (2024). *Fus-MAE: A cross-attention-based data fusion approach for Masked Autoencoders in remote sensing*. arXiv (Cornell University).

Chenari, B., Lamas, F. B., Gaspar, A. R., & Silva, M. G. d. (2017). Development of a new CO2-based demand-controlled ventilation strategy using EnergyPlus. https://doi.org/10.31224/osf.io/h6mdn.

Cincinelli, A., & Martellini, T. (2017). *Indoor air quality and health*. International Journal of Environmental Research and Public Health, 14(11), 1286. https://doi.org/10.3390/ ijerph14111286

Cipollone, V., Morresi, N., Serroni, S., Casaccia, S., Giovanardi, M., Pracucci, A., Arnone, D., & Revel, G. M. (2024). *AI-based methodology for thermal comfort measurement: Application of a simplified comfort model on a real-life case study*. 2024 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv), 11, 286–291. https://doi.org/10.1109/metrolivenv60384.2024.10615461

Cretescu, I., Isopescu, D. N., Lutic, D., & Soreanu, G. (2019). *Indoor air pollutants and the future perspectives for living space design*. Indoor Environment and Health. https://doi.org/10.5772/intechopen.87309

Deng, Z., Dong, B., Guo, X., & Zhang, A. (2023). *A pilot study on the combined multi-domain impact of indoor air quality and noise on office productivity*. E3S Web of Conferences, 396, 01047. https://doi.org/10.1051/e3sconf/202339601047

Du, Y., Zandi, H., Kotevska, O., Kurte, K., Munk, J., Amasyali, K., McKee, E., & Li, F. (2021). *Intelligent multi-zone residential HVAC control strategy based on deep reinforcement learning*.

Applied Energy, 281, 116117. https://doi.org/10.1016/j.apenergy.2020.116117

EPA. (2023). *Indoor air quality*. Accessed October 12, 2024. https://www.epa.gov/report-environment/ indoor-air-quality

Fang, Z., Crimier, N., Scanu, L., Midelet, A., Alyafi, A., & Delinchant, B. (2021). *Multi-zone indoor temperature prediction with LSTM-based sequence to sequence model*. Energy and Buildings, 245, 111053. https://doi.org/10.1016/j.enbuild.2021.111053

Fritz, H., Tang, M., Kinney, K., & Nagy, Z. (2022). *Evaluating machine learning models to classify occupants' perceptions of their indoor environment and sleep quality from indoor air quality*. Journal of the Air & Waste Management Association. https://doi.org/10.1080/10962247. 2022.2105439

Gao, J., Li, P., Chen, Z., & Zhang, J. (2020). *A survey on deep learning for multimodal data fusion*. Neural Computation, 32(5), 829–864. https://doi.org/10.1162/neco_a_01273

Jones, A. P. (1999). *Indoor air quality and health*. Atmospheric Environment, 33(28), 4535–4564. https://doi.org/10.1016/s1352-2310(99)00272-1

Kapoor, N. R., Kumar, A., Kumar, A., Kumar, A., Mohammed, M. A., Kumar, K., Kadry, S., & Lim, S. (2022). *Machine learning-based CO_2 prediction for office room: A pilot study*. Wireless Communications and Mobile Computing, 2022, 1–16. https://doi.org/10.1155/2022/9404807

Lahat, D., Adali, T., & Jutten, C. (2015). *Multimodal data fusion: An overview of methods, challenges, and prospects*. Proceedings of the IEEE, 103(9), 1449–1477. https://doi.org/10.1109/jproc.2015.2460697

Lee, M. J., & Zhang, R. (2024). *Human-centric artificial intelligence of things–based indoor environment quality modeling framework for supporting student well-being in educational facilities*. Journal of Computing in Civil Engineering, 38(2). https://doi.org/10.1061/jccee5.cpeng-5632

Lin, X., Chao, S., Yan, D., Guo, L., Liu, Y., & Li, L. (2023). *Multi-sensor data fusion method based on self-attention mechanism*. Applied Sciences, 13(21), 11992. https://doi.org/10.3390/ app132111992

Liu, J., Capurro, D., Nguyen, A., & Verspoor, K. (2023). *Attention-based multimodal fusion with contrast for robust clinical prediction in the face of missing modalities*. Journal of Biomedical Informatics, 145, 104466. https://doi.org/10.1016/j.jbi.2023.104466

Liu, X., Ren, M., Yang, Z., Yan, G., Guo, Y., Cheng, L., & Wu, C. (2022). *A multi-step predictive deep reinforcement learning algorithm for HVAC control systems in smart buildings*. Energy, 124857. https://doi.org/10.1016/j.energy.2022.124857

Lolli, F., Marinello, S., Coruzzolo, A., & Butturi, M. (2022). *Post-occupancy evaluation's (POE) applications for improving indoor environment quality (IEQ)*. Toxics, 10(10), 626. https: //doi.org/10.3390/toxics10100626

Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). *A survey on machine learning for data fusion*. Information Fusion, 57, 115–129. https://doi.org/10.1016/j.inffus.2019.12.001

Mitkov, R., Petrova-Antonova, D., & Hristov, P. O. (2023). *Predictive modeling of indoor environmental parameters for assessing comfort conditions in a kindergarten setting*. Toxics, 11(8), 709. https://doi.org/10.3390/toxics11080709

Rizk, H., Torki, M., & Youssef, M. (2019). *Cellindeep: Robust and accurate cellular-based indoor localization via deep learning*. IEEE Sensors Journal, 19(6), 2305–2312. https://doi.org/ 10.1109/jsen.2018.2885958

Samet, J. M., & Spengler, J. D. (2003). *Indoor environments and health: Moving into the 21st century*. American Journal of Public Health, 93(9), 1489–1493. https://doi.org/10.2105/ajph.93. 9.1489

Shamsudin, S. (2023). *Indoor air quality assessment in the office of the transformer manufacturing factory in Selangor, Malaysia*. Malaysian Journal of Medicine and Health Sciences, 19(5), 24–31. https://doi.org/10.47836/mjmhs.19.5.5

Shan, X., Melina, A. N., & Yang, E.-H. (2018). *Impact of indoor environmental quality on students' wellbeing and performance in educational building through life cycle costing perspective*. Journal of Cleaner Production, 204, 298–309. https://doi.org/10.1016/j.jclepro.2018.09. 002

Smarra, F., Jain, A., de Rubeis, T., Ambrosini, D., D'Innocenzo, A., & Mangharam, R. (2018). *Datadriven model predictive control using random forests for building energy optimization and climate control*. Applied Energy, 226, 1252–1272. https://doi.org/10.1016/j.apenergy. 2018.02.126

Sun, S., Zheng, X., Villalba-D´ıez, J., & Ordieres-Mer´e, J. (2019). *Indoor air-quality data-monitoring system: Long-term monitoring benefits*. Sensors, 19(19), 4157. https://doi.org/10.3390/ s19194157

Tagliabue, L. C., Re Cecconi, F., Rinaldi, S., & Ciribini, A. L. C. (2021). *Data-driven indoor air quality prediction in educational facilities based on IoT network*. Energy and Buildings, 236, 110782. https://doi.org/10.1016/j.enbuild.2021.110782

Tien, P., Wei, S., Darkwa, J., Wood, C., & Calautit, J. (2022). *Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality – A review*. Energy and AI, 10, 100198. https://doi.org/10.1016/j.egyai.2022.100198

USGBC. (2014). *Green Building 101: What is indoor environmental quality?* U.S. Green Building Council. https://www.usgbc.org/articles/green-building-101-what-indoor-environmental-quality

Vardoulakis, S., Giagloglou, E., Steinle, S., Davis, A., Sleeuwenhoek, A., Galea, K., ... & Crawford, J. (2020). *Indoor exposure to selected air pollutants in the home environment: A systematic review*. International Journal of Environmental Research and Public Health, 17(23), 8972. https://doi.org/10.3390/ijerph17238972

Xing, T., Sun, K., & Zhao, Q. (2023). *MITP-Net: A deep-learning framework for short-term indoor temperature predictions in multi-zone buildings*. Building and Environment, 239, 110388. https://doi.org/10.1016/j.buildenv.2023.110388

Zahaba, M., Tamsi, N., Engliman, S., Kamaruddin, A., Hassan, N., & Ariffin, A. (2022). *Indoor air quality (IAQ) in a naval ship after refit program: A time variation analysis*. IOP Conference Series: Earth and Environmental Science, 1013(1), 012004. https://doi.org/10.1088/ 1755-1315/1013/1/012004

Zhang, X., & Li, P. (2023). *Transfer learning in the Transformer model for thermal comfort prediction: A case of limited data*. Energies, 16(20), 7137. https://doi.org/10.3390/en16207137

Zhao, Y., Guo, S., Chen, Z., Shen, Q., Meng, Z., & Xu, H. (2022). *MARFusion: An attentionbased multimodal fusion model for human activity recognition in real-world scenarios*. Applied Sciences, 12(11), 5408. https://doi.org/10.3390/app12115408

Zhu, Y., Al-Ahmed, S. A., Shakir, M. Z., & Olszewska, J. I. (2022). *LSTM-based IoT-enabled CO_2 steady-state forecasting for indoor air quality monitoring*. Electronics, 12(1), 107. https: //doi.org/10.3390/electronics12010107