AI-Driven Indoor Air Quality Prediction and Health Risk Assessment of Volatile Organic Compounds

Danqing Duan
Department of Computer Science, University of New Hampshire, Durham, NH, USA.
danqingwork@unh.edu

Abstract

Indoor air quality is a persistent public health challenge because volatile organic compounds emitted from building materials, furnishings, cleaning products, and occupant activities can produce chronic and acute health effects that are not well captured by episodic monitoring. The increasing availability of low-cost sensors, building management data, and machine learning methods has created new opportunities for continuous indoor air quality prediction. However, the translation of predicted volatile organic compound concentrations into meaningful health risk information requires more than model accuracy. It demands careful attention to system architecture, exposure science, data governance, uncertainty communication, robustness, fairness, and deployment context. This paper presents a system-level analysis of AI-driven indoor air quality prediction and health risk assessment for volatile organic compounds. It examines the structural design of sensing and modeling pipelines, the integration of toxicological and exposure considerations, the governance of occupant-related data, and the policy trade-offs that arise when predictive systems are deployed in heterogeneous buildings. The discussion emphasizes that AI-based prediction should be understood as part of a larger socio-technical infrastructure rather than as an isolated computational task. A well-designed system must balance predictive performance with interpretability, privacy protection, equity, and operational sustainability. The paper concludes by identifying directions for interdisciplinary research and policy development to ensure that AI-enhanced indoor air quality management improves health outcomes without exacerbating existing environmental disparities.

Keywords

indoor air quality; volatile organic compounds; artificial intelligence; health risk assessment; smart buildings; data governance.

1. Introduction

Indoor air quality has become a central concern in public health and building science because urban populations spend the majority of their time in enclosed environments where pollutant concentrations can differ substantially from outdoor levels. International guidance has identified selected indoor pollutants as priorities for policy and risk management [1]. Early systematic assessments established that indoor air pollution contributes to both acute and chronic disease burdens, with effects ranging from sensory irritation to respiratory and neurological outcomes [2]. Over several decades, changes in building materials, ventilation practices, and consumer products have altered the mixture of pollutants encountered indoors, producing new exposure patterns that are not fully addressed by traditional regulatory frameworks [3].

Among indoor pollutants, volatile organic compounds are especially challenging because they include hundreds of chemical species emitted from building materials, furnishings, cleaning agents, and occupant activities. Chronic health impact estimation for organic compounds in residences has shown that a relatively small number of compounds can drive much of the noncancer and cancer risk, yet exposure remains highly variable across buildings and populations [4]. Formaldehyde, as one of the most studied indoor carbonyls, illustrates the difficulty of controlling emissions from composite wood products, coatings, and other durable sources even when ventilation standards are met [5]. Monitoring campaigns across European dwellings have further demonstrated that measured concentrations of benzene, toluene, ethylbenzene, xylenes, formaldehyde, and acetaldehyde vary by orders of magnitude depending on building age, season, and occupant behavior [6].

Although traditional sampling and laboratory analysis remain essential, they cannot capture the continuous spatiotemporal dynamics needed for proactive health protection. Artificial intelligence offers the possibility of inferring pollutant concentrations from lower-cost sensors, building operational data, and contextual variables. However, the transition from predictive accuracy to trustworthy health risk assessment requires system-level attention to architecture, data governance, uncertainty, fairness, robustness, and deployment context. This paper examines these structural dimensions and proposes an integrated conceptual framework for AI-driven indoor air quality prediction and health risk assessment of volatile organic compounds.

2. Volatile Organic Compound Exposure and Health Risk Complexity

The assessment of volatile organic compound exposure begins with human time-activity patterns. National human activity surveys have repeatedly found that most individuals spend the large majority of each day in residential and nonresidential indoor environments, making indoor air the dominant inhalation exposure medium for many volatile compounds [7]. Longitudinal reviews of indoor volatile organic compound concentrations have documented substantial variability across buildings and seasons, even when comparative studies use similar sampling protocols [8]. In addition to direct emissions, indoor chemistry can create secondary pollutants. For example, reactions between unsaturated terpenes from cleaning products and ozone can generate aldehydes, ultrafine particles, and other species that are not emitted directly from a single source [9].

The health risk associated with volatile organic compounds is not a simple function of average concentration. Risk depends on the duration of exposure, the toxicological potency of each compound, the presence of co-occurring stressors, and the susceptibility of occupants. Residential environments are particularly important because exposures can be prolonged and include sensitive groups such as children, older adults, and people with pre-existing respiratory conditions. Building materials can release pollutants over months or years, and emission rates respond to temperature, humidity, and ventilation. These dynamic processes create a mismatch between episodic measurement campaigns and the temporal granularity needed for health risk management. Artificial intelligence-driven prediction is therefore attractive not because it replaces environmental chemistry, but because it can provide continuous estimates that integrate multiple determinants of exposure.

3. AI-Driven Indoor Air Quality Prediction Architecture

A system-level artificial intelligence architecture for indoor volatile organic compound prediction typically consists of a sensing layer, a data ingestion and feature engineering layer, a modeling layer, and an output layer that supports decision making. Reviews of machine learning applied to indoor air quality indicate that methods such as random forests, support vector machines, gradient boosting, and neural networks have been used to predict concentrations of carbon dioxide, particulate matter, and selected gaseous pollutants [10]. Gradient boosting approaches have proven effective for tabular building data because they capture nonlinear interactions and are relatively robust to missing values [11]. For temporal sequences, long short-term memory networks provide a mechanism to retain information over time, which is valuable when pollutant concentrations respond to lagged ventilation and emission events [12]. More recent sequence modeling architectures based on self-attention can capture long-range dependencies in building operational data, although they require larger datasets and careful regularization [13]. Deep learning more broadly has expanded the capacity to learn hierarchical representations from raw sensor streams, but it also introduces concerns about data requirements and interpretability [14].

Model interpretation is not only a post hoc exercise but should be embedded into the architecture. Local explanation methods can reveal which input features drive a prediction for a particular building or time window, helping facility managers distinguish sensor drift from genuine pollution events [15]. Occupant behavior is one of the most influential and least predictable inputs in indoor environmental modeling, and building-level models must account for window opening, cooking, cleaning, use of personal care products, and occupancy patterns [16]. These behavioral variables are difficult to collect continuously without violating privacy or imposing high sensor burdens, so the prediction system often relies on proxies such as door sensors, motion detectors, energy use, and indoor concentration signals. The architecture therefore requires a trade-off between high-resolution behavioral monitoring and the practical limits of occupancy sensing. A layered design can address this by separating physics-informed source models, data-driven black-box models, and domain knowledge modules that constrain predictions to plausible ranges.

4. Health Risk Assessment Integration and Uncertainty

The translation of predicted volatile organic compound concentrations into health risk estimates requires a coherent exposure and toxicological framework. A recent residential volatile organic compound study has quantified exposures from building materials, associated health risks, and ambient hazards, reinforcing the importance of source-specific characterization [17]. That work also illustrates how risk can be distributed unequally across residential archetypes, ventilation rates, and material choices. The integration of artificial intelligence predictions with health risk assessment should distinguish between screening-level estimates used for building management and more rigorous regulatory assessments that require compound-specific toxicity values and long-term exposure distributions. Indoor aerosol and gas exposure studies further emphasize that personal exposure cannot be inferred from a single room sensor without accounting for spatial inhomogeneity and particle dynamics [18]. Recent conceptual work on occupant health in buildings during normal operation and extreme events has similarly emphasized that indoor air quality management must account for vulnerable populations, mental health, and the combined effects of thermal, acoustic, and chemical stressors [19]. Consequently, an artificial intelligence-based health risk layer should not merely report concentrations but should produce risk categories, confidence intervals, and context-specific warnings.

Uncertainty is structural in this pipeline. It arises from sensor error, missing data imputation, model approximation, exposure factor variability, and toxicological uncertainty. A robust risk assessment system should propagate these uncertainties and communicate them in ways that support actionable decisions. For example, a predicted short-term increase in formaldehyde may be more uncertain than an elevated baseline associated with low ventilation, and the health risk interpretation should reflect that distinction. The system should also avoid deterministic overconfidence. Presenting a single numerical risk estimate may create an illusion of precision that undermines trust when the estimate changes under modest input variation. Instead, continuous Bayesian or ensemble approaches can provide ranges that help building operators distinguish between routine fluctuations and meaningful intervention thresholds.

5. Data Infrastructure, Governance, and Interoperability

The data foundation for AI-driven volatile organic compound risk assessment includes environmental sensors, building management systems, weather feeds, occupancy information, and material inventories. Low-cost monitoring systems have expanded the availability of indoor environmental data, but they also raise questions about calibration, long-term drift, and data comparability [20]. The architecture should include automated quality control procedures, periodic reference calibration, and metadata standards that preserve information about sensor type, location, sampling interval, and known interference. Without such infrastructure, models trained in one building may produce unreliable outputs when transferred to another building with different sensor characteristics or ventilation configurations. Interoperability is particularly important because indoor air quality data must be linked with building information models, energy management systems, and health-relevant contextual data.

Governance is equally important because prediction and risk assessment depend on data that may reveal sensitive information about occupancy,Governance is equally important because prediction and risk assessment depend on data that may reveal sensitive information about occupancy, activity patterns, health status, and daily routines. This issue is amplified by the fact that indoor environmental data are frequently collected in private spaces such as homes, rental apartments, workplaces, schools, and healthcare facilities. The information generated by sensors and models can be used to infer whether a dwelling is occupied, when occupants wake and sleep, whether they cook or clean at particular times, and whether they may have respiratory conditions that increase sensitivity to pollutants. These inferences are not always explicit in the raw data, but they emerge through the combination of multiple data streams and the use of machine learning models. Privacy risks therefore cannot be addressed solely by anonymizing sensor identifiers; they require a broader governance framework that limits data linkage, restricts secondary uses, and gives occupants meaningful control over how their environmental data are used.

Data protection obligations may vary across jurisdictions, but the underlying principle of proportionality remains relevant. An AI-based indoor air quality system should collect only the data that are necessary to achieve the stated public health or building management objective. For example, a model that predicts formaldehyde concentrations does not require continuous video surveillance or high-resolution location tracking. Occupancy can be inferred from carbon dioxide sensors, door contact sensors, or coarse motion detectors, and model performance may be acceptable without continuous identification of individuals. The governance architecture must therefore distinguish between technical feasibility and operational necessity. It must also define how long data are retained, who may access model predictions, and how risk outputs are communicated to occupants. Without such safeguards, building managers may inadvertently create detailed behavioral profiles that could be used for purposes unrelated to health protection, including tenant evaluation, insurance pricing, or workplace surveillance.

Interoperability is a further structural concern. Many buildings already contain separate systems for heating, ventilation, air conditioning, lighting, access control, and energy metering. These systems are often procured from different vendors and use incompatible data formats. An AI-driven indoor air quality prediction service cannot operate in isolation from this broader building information environment. It must be able to integrate data from building management systems, occupant feedback platforms, weather services, and material inventories while preserving semantic consistency and temporal alignment. Standardized metadata descriptions are essential for long-term operation. Without them, a model trained in a mechanically ventilated office building may be inappropriately applied to a naturally ventilated dwelling with different source profiles and occupant behaviors. Governance and technical standardization are therefore closely linked: data sharing protocols, application programming interfaces, and model documentation all shape the portability and trustworthiness of predictive systems across heterogeneous building stocks.

6. Structural Trade-offs and Deployment Considerations

The deployment of AI-driven volatile organic compound prediction involves a series of structural trade-offs that cannot be resolved solely by optimizing a single performance metric. One central trade-off concerns model complexity and interpretability. Highly flexible models such as deep neural networks can capture nonlinear interactions and temporal dependencies, but they may require large amounts of training data and may produce outputs that are difficult for building managers to interpret. Simpler models such as gradient boosting or regularized regression may be easier to explain and maintain, but they can miss complex emission dynamics or time-lagged effects associated with ventilation changes. The choice of model family should be guided by the operational context, the available data, the expected health decision, and the capacity of the responsible organization to monitor model performance over time. A system designed for continuous health risk screening in residential buildings may benefit from a modular architecture in which physics-based source models constrain the outputs of data-driven models, thereby reducing the risk of biologically implausible predictions.

A second structural trade-off involves the distribution of computation between local and centralized resources. Edge computation can reduce the transmission of sensitive data and improve responsiveness, but it may limit the complexity of models that can be executed on low-power sensor nodes. Cloud-based inference can support more sophisticated models and centralized monitoring, but it raises data governance risks and depends on reliable connectivity. Hybrid architectures that perform preliminary quality control at the edge and more complex risk inference in a centralized environment can balance these concerns, provided that the data transmitted to the cloud are minimized and protected. The operational sustainability of the system also depends on calibration and maintenance. Low-cost sensors can drift over time, especially in humid or polluted indoor environments, and a model trained at installation may become unreliable if sensors are not periodically recalibrated. Deployment plans should therefore include automated drift detection, reference calibration protocols, and clear thresholds for model retraining.

The relationship between ventilation and energy consumption is another important deployment consideration. Smart ventilation systems can respond to predicted pollutant levels to improve air quality, but they can also increase heating and cooling loads if ventilation rates are raised during extreme outdoor conditions. Previous work on smart ventilation has explored the trade-off between energy use and indoor air quality in residential buildings, showing that demand-controlled ventilation can improve air quality without proportional increases in energy consumption when controllers are well designed [21]. This finding has direct implications for AI-based prediction because a model that accurately predicts volatile organic compound concentrations can be used to optimize ventilation only if the system also accounts for energy costs, thermal comfort, and equipment constraints. The health benefit of increased ventilation must be weighed against the potential for higher carbon emissions, especially in buildings that rely on fossil-fuel-based heating or cooling. This concern is amplified by the fact that performance disparities across housing types can arise from uneven sensor availability and data coverage, a problem closely related to algorithmic fairness [22].

Ventilation is itself a determinant of health that interacts with volatile organic compound concentrations. A multidisciplinary review of the scientific literature has documented associations between ventilation rates and health outcomes, particularly respiratory symptoms and infectious disease transmission [24]. This body of evidence supports the integration of ventilation rate predictions into health risk assessments, but it also illustrates the complexity of causal interpretation. Occupants may experience health effects from airborne pollutants, inadequate thermal conditions, or aeroallergens, and these factors can be correlated. AI systems that predict volatile organic compound levels should therefore be designed as part of a broader indoor environmental quality framework rather than as a standalone pollutant model. A risk assessment that ignores ventilation, temperature, humidity, and occupancy may produce misleading recommendations, particularly in diverse residential settings where building characteristics vary widely.

7. Fairness, Robustness, and Equity

Fairness is a structural property of the entire prediction and risk assessment pipeline, not merely a post hoc adjustment to model outputs. Machine learning models trained on data from high-end buildings with dense sensor networks may perform poorly when transferred to older buildings with fewer sensors, different construction materials, and higher occupant density. The resulting performance disparities can produce unequal health protection if model errors lead to missed interventions in disadvantaged housing. The growing literature on bias and fairness in machine learning emphasizes that fairness cannot be defined in a single universal way, because different definitions may conflict and because the social context determines which harms are most relevant [22]. In the indoor air quality domain, the most pressing fairness concern is not statistical parity across demographic groups but rather the equitable distribution of reliable predictions and actionable risk information across building types and socioeconomic contexts. A model that accurately predicts volatile organic compound concentrations in newly constructed commercial buildings but fails in public housing is not merely less accurate; it may reinforce existing environmental health disparities.

Robustness is closely related to fairness. Predictive models are vulnerable to sensor drift, missing data, concept drift, and adversarial manipulation. Sensor drift can introduce systematic bias that is not captured by standard validation metrics if the validation data are collected under ideal conditions. Missing data are common in real buildings because sensors are temporarily disconnected, batteries fail, or network outages interrupt data transmission. Models that perform well on complete datasets may degrade rapidly when faced with realistic missingness patterns. Concept drift occurs when the relationship between measured variables and volatile organic compound concentrations changes because of building renovation, new furniture, changed cleaning products, or altered occupancy patterns. Robust deployment requires continuous monitoring of model inputs and outputs, automatic detection of distribution shifts, and a plan for retraining or recalibration. It also requires careful treatment of uncertainty so that users do not treat a changing risk estimate as a system failure when in fact it reflects genuine environmental variability.

The governance of algorithmic systems must address the ethical and political dimensions of automation. The ethics of algorithms has developed into a substantial field concerned with transparency, accountability, and the distribution of power in data-driven systems [23]. In the context of indoor air quality, algorithmic governance involves decisions about who defines acceptable risk thresholds, who is notified when predictions exceed those thresholds, and who bears responsibility when a model fails to warn occupants of elevated volatile organic compound concentrations. These are not purely technical questions. They involve labor relations, tenant rights, public health authority, and the obligations of building owners. A building manager may prefer to receive only a binary alert, while a public health agency may require probabilistic estimates and model documentation. Occupants may want explanations that are comprehensible and actionable, such as a recommendation to increase ventilation or to avoid using certain cleaning products. Designing governance structures that accommodate these different needs is essential for building trust and ensuring that predictive systems are used in ways that protect health without undermining autonomy.

8. Policy and Regulatory Implications

The integration of AI-driven prediction into indoor air quality policy requires an evolving regulatory framework that can accommodate both the dynamic nature of machine learning systems and the established principles of environmental health protection. Existing building codes and ventilation standards focus primarily on minimum ventilation rates and source control measures, but they do not yet address the continuous prediction of volatile organic compound concentrations using data-driven methods. Policymakers should consider establishing performance-based criteria for predictive systems rather than prescribing specific model architectures. Such criteria might include validation requirements under representative indoor conditions, documentation of sensor accuracy and calibration procedures, and transparent reporting of uncertainty. Certification schemes for low-cost sensors and machine learning services could help building owners distinguish between robust systems and commercial products that promise more than they can deliver.

Health-based benchmarks are also important. A useful AI risk layer should translate concentration predictions into risk categories that correspond to existing exposure guidelines or health effect thresholds. Controlled exposure studies have demonstrated associations between indoor volatile organic compounds, carbon dioxide, ventilation, and cognitive function in office workers, suggesting that even moderate indoor environmental changes can influence performance and well-being [25]. These findings support the inclusion of cognitive and neurological outcomes in health risk assessments, but they also complicate the regulatory picture because risk communication must avoid overstating temporary performance effects as long-term disease risks. A policy framework should therefore distinguish between short-term acute irritation, longer-term chronic risks, and functional outcomes such as cognitive performance, and should require that AI-generated risk information be accompanied by clear explanations of the underlying evidence and its limitations.

Finally, policy interventions should address the economic incentives that shape deployment. Building owners may be reluctant to invest in continuous monitoring and predictive risk systems if the benefits accrue primarily to occupants or public health authorities. Subsidies, tax incentives, or inclusion in green building certification programs can align private and public interests. Data sharing policies can also encourage the development of robust models while protecting privacy through aggregation, federated learning, and strict access controls. However, data sharing must not become a new burden for low-income building owners who lack technical capacity. Equity-oriented policy should ensure that AI-based indoor air quality tools are not only available to high-end commercial buildings but are also adapted to public housing, schools, childcare centers, and healthcare facilities. This requires deliberate investment in reference datasets that represent diverse building stocks, as well as open benchmarks that enable independent evaluation.

9. Conclusion

AI-driven prediction of volatile organic compounds and associated health risk assessment represents a promising but structurally demanding integration of environmental sensing, machine learning, exposure science, and public health governance. The value of these systems lies not in the sophistication of any single model but in the coherence of the entire pipeline, from sensor deployment and data governance to uncertainty communication and policy integration. Volatile organic compounds pose distinctive challenges because they arise from multiple indoor sources, respond dynamically to ventilation and occupant activity, and include compounds with different toxicological profiles. A predictive system that treats all volatile organic compounds as a single aggregate may be useful for ventilation control but is insufficient for health risk assessment. A more rigorous approach requires compound-specific predictions, exposure factor characterization, and transparent uncertainty bounds.

The structural trade-offs identified in this paper show that there is no universal optimal architecture for AI-based indoor air quality prediction. The appropriate balance between model complexity and interpretability, local and centralized computation, privacy and monitoring granularity, and energy use and ventilation depends on the characteristics of the building stock, the population served, and the health decisions to be supported. Robustness and fairness must be evaluated across heterogeneous deployment contexts rather than in a single laboratory or demonstration building. Algorithmic fairness in this domain should focus on equitable access to reliable risk information and on preventing model failures in vulnerable housing. Governance arrangements must ensure that data collected for health protection are not repurposed in ways that harm occupants or undermine trust.

Future research should develop open benchmark datasets that represent residential and nonresidential buildings across different climates, construction eras, and socioeconomic settings. Interdisciplinary collaboration between building scientists, toxicologists, computer scientists, and social scientists is necessary to translate model outputs into health-relevant interventions. The policy environment must evolve in parallel, establishing clear expectations for system validation, risk communication, and liability. Only through such a system-level approach can AI-driven indoor air quality prediction fulfill its potential to reduce volatile organic compound exposure and improve public health without reinforcing existing environmental disparities.

References

1. World Health Organization. (2010). WHO guidelines for indoor air quality: Selected pollutants. WHO Regional Office for Europe.

2. Samet, J. M., Marbury, M. C., & Spengler, J. D. (1987). Health effects and sources of indoor air pollution. Part I. American Review of Respiratory Disease, 136(6), 1486–1508.

3. Weschler, C. J. (2009). Changes in indoor pollutants since the 1950s. Atmospheric Environment, 43(1), 153–169.

4. Logue, J. M., McKone, T. E., Sherman, M. H., & Singer, B. C. (2011). Hazard assessment of chemical air contaminants measured in residences. Indoor Air, 21(2), 92–109.

5. Salthammer, T., Mentese, S., & Marutzky, R. (2010). Formaldehyde in the indoor environment. Chemical Reviews, 110(4), 2536–2572.

6. Sarigiannis, D. A., Karakitsios, S. P., Gotti, A., Liakos, I. L., & Katsoyiannis, A. (2011). Exposure to major volatile organic compounds and carbonyls in European indoor environments and associated health risk. Environment International, 37(4), 743–765.

7. Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., Behar, J. V., Hern, S. C., & Engelmann, W. H. (2001). The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants. Journal of Exposure Analysis and Environmental Epidemiology, 11(3), 231–252.

8. Brown, S. K., Sim, M. R., Abramson, M. J., & Gray, C. N. (1994). Concentrations of volatile organic compounds in indoor air – A review. Indoor Air, 4(2), 123–134.

9. Weschler, C. J., & Shields, H. C. (1999). Indoor ozone/terpene reactions as a source of indoor particles. Atmospheric Environment, 33(15), 2301–2312.

10. Wei, W., Ramalho, O., Mandin, C., & Cattaneo, A. (2015). Indoor air quality modeling using machine learning: A review. Building and Environment, 94, 460–469.

11. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM.

12. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

13. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems 30 (pp. 5998–6008). Curran Associates.

14. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

15. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). ACM.

16. Yan, D., O'Brien, W., Hong, T., Feng, X., Burak Gunay, H., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. Energy and Buildings, 94, 264–278.

17. Fan, Yi, et al. "Residential VOC from building materials: Exposures, health risks, and ambient hazards." Sustainable Cities and Society 119 (2025): 106080.

18. Nazaroff, W. W. (2004). Indoor particle dynamics. Indoor Air, 14(Suppl 7), 175–183.

19. Bluyssen, P. M. (2014). The healthy indoor environment: How to assess occupants' wellbeing in buildings. Routledge.

20. Kumar, P., Skouloudis, A. N., Bell, M., Viana, M., Carotta, M. C., Biskos, G., & Morawska, L. (2016). Real-time sensors for indoor air monitoring and challenges ahead in deploying them to urban areas. Science of the Total Environment, 560-561, 150–159.

21. Sherman, M. H., & Walker, I. S. (2011). Meeting residential ventilation standards through dynamic control of ventilation systems. Energy and Buildings, 43(8), 1904–1912.

22. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning: Limitations and opportunities. fairmlbook.org.

23. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.

24. Sundell, J., Levin, H., Nazaroff, W. W., Cain, W. S., Fisk, W. J., Grimsrud, D. T., Gyntelberg, F., Li, Y., Persily, A. K., Pickering, A. C., Samet, J. M., Spengler, J. D., Taylor, S. T., & Weschler, C. J. (2011). Ventilation rates and health: Multidisciplinary review of the scientific literature. Indoor Air, 21(3), 191–204.

25. Allen, J. G., MacNaughton, P., Satish, U., Santanam, S., Vallarino, J., & Spengler, J. D. (2016). Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: A controlled exposure study of green and conventional office environments. Environmental Health Perspectives, 124(6), 805–812.