

# Temporal and Spatial Dynamics of Public Sentiment Toward Indoor Ventilation in New York City: Evidence From Social Sensing and Large Language Models

Mehdi Ashayeri<sup>1</sup>

Southern Illinois University Carbondale, Illinois, United States

## Abstract:

This study examines public sentiment toward indoor ventilation in New York City using social sensing and large language models applied to geotagged Twitter data from 2013 to 2023. A longitudinal corpus of 1,500 tweets was compiled and analyzed to identify temporal and spatial patterns in ventilation-related discourse across New York City boroughs. After data cleaning and relevance screening, sentiment classification was conducted using RoBERTa to distinguish positive and negative expressions associated with indoor ventilation. The results show that discussion of indoor ventilation remained limited before 2019 but increased sharply during the COVID-19 period, with the highest volume observed in 2020 and sustained activity through 2022. Spatial analysis further indicates that Manhattan accounted for the largest concentration of ventilation-related tweets, followed by Brooklyn and Queens, suggesting uneven borough-level engagement with indoor air issues. Model confidence scores remained relatively stable across years, supporting the consistency of the classification results. Overall, the findings demonstrate how Twitter, as a form of social sensing, can help reveal changing public perceptions of indoor ventilation in urban environments, while large language models provide an effective approach for interpreting sentiment in geolocated public health discourse.

**Keywords:** ventilation; social sensing; large language models; X/Twitter; public health

## 1. Introduction

As metropolitan centers continue to densify, the tension between high-occupancy environments and the urgent necessity for healthy indoor infrastructure has become a critical focal point for public safety (Hartmann, 2022). Current building performance standards often prioritize operational energy efficiency over the physical well-being of the workforce and residents (Perera et al., 2025). This institutional oversight is particularly concerning given that modern urban dwellers spend more than ninety percent of their lives in indoor spaces where air quality is rarely monitored or optimized (Allen et al., 2016). Consequently, the integration of advanced sensing technologies and data-driven analytics is now essential to bridge the disconnect between architectural engineering and healthy public health outcomes (Shi et al., 2021).

Building ventilation serves as a primary defense mechanism in urban public health by regulating air exchange rates to dilute and remove airborne pathogens (EPA, 2025; Morawska et al., 2020). However, increasing mechanical ventilation is a complex environmental intervention; rapid air exchange in multi-room buildings can occasionally facilitate the spread of viral particles to adjacent rooms by elevating concentrations faster than they can be filtered (Pease et al., 2021). Leading regulatory bodies have recently updated their technical guidelines to recommend a

minimum of five air changes per hour (ACH) to significantly mitigate respiratory transmission risks (California Department of Public Health, 2021). Despite these engineering standards, the practical implementation of high-efficiency filtration remains highly inconsistent across the existing diverse urban building stock (Global Virus Network, 2023). Therefore, the post-pandemic recovery requires a fundamental re-evaluation of how existing mechanical systems are maintained to ensure occupant safety (Morawska et al., 2020).

Social sensing has emerged as a novel paradigm in urban studies, utilizing crowd-sourced data streams to capture real-time reflections of the city's functional state and resident sentiment (Kamel Boulos et al., 2011; Shi et al., 2021). These user-generated digital footprints offer a high-resolution perspective on urban experiences that traditional stationary physical sensors cannot fully encapsulate (Son & Thill, 2018). By leveraging decentralized platforms such as Twitter and Quora, researchers can identify emerging environmental concerns and public health anxieties with unprecedented temporal sensitivity (Hong et al., 2022). This approach effectively transforms city residents into a distributed sensor network, providing unsolicited insights into the changing priorities of metropolitan populations (Epifantsev et al., 2024). Such crowd-sourced data sources are increasingly vital for urban planners seeking to align infrastructure retrofitting with the actual subjective needs and perceptions of the public (Hong et al., 2022).

The COVID-19 pandemic acted as a unique and profound temporal shock that significantly heightened public awareness regarding indoor air quality and environmental safety (Rashid & Wang, 2021; Rousseau & Deschacht, 2020). During the acute phases of the crisis, social media platforms served as essential channels for individuals to report their localized observations and health experiences (Rashid & Wang, 2021). This period also saw a noticeable shift in public interest toward environmental health and nature-related topics as lockdown measures restricted movement and emphasized the value of ventilated spaces (Rousseau & Deschacht, 2020). Furthermore, the urgency of the health emergency catalyzed a qualitative change in how residents perceive the resilience and safety of their local indoor environments (Hong et al., 2022). This evolving public sentiment provides a critical research window for studying the long-term impact of the pandemic on urban infrastructure expectations (Bertram & Rehdanz, 2015; Samuelsson et al., 2018).

The advent of Large Language Models (LLMs) has revolutionized the capacity of natural language processing (NLP) to handle the complex, unstructured social sentiment inherent in urban data (Jiang et al., 2025; Li et al., 2025). Transformer-based architectures, such as BERT, enable the automated extraction of nuanced semantic patterns and sentiment from massive collections of citizen feedback (Devlin et al., 2019; Ho et al., 2024). These models significantly outperform traditional machine learning techniques by effectively capturing long-range dependencies and bidirectional context in human expression (Zhang et al., 2024). In the context of urban informatics, LLMs facilitate the integration of disparate textual data sources to provide a richer, more actionable understanding of public opinion (Jiang et al., 2025). By automating the coding of qualitative sentiment at scale, these advanced AI systems empower urban policymakers to derive evidence-based insights for city-wide interventions (Ho et al., 2024).

Human perception of indoor health frequently diverges from objective environmental engineering standards, necessitating a more holistic approach to resident well-being (ASHRAE, 2025; Dengel,

2024). Subjective feelings of safety and comfort are influenced by complex interactions between ventilation, lighting, and acoustic parameters within the built environment (Hernandez-Martin et al., 2025). While established frameworks like the WELL Building Standard provide a roadmap for healthy spaces, the actual lived experience of occupants is often overlooked in building retrofits (Perera et al., 2025). Poor indoor air quality has been directly linked to neurological outcomes such as reduced productivity and impaired cognitive function in workplace settings (Morawska et al., 2021). Ensuring that building retrofits satisfy both physical health metrics and psychological human perceptions of safety is therefore a critical frontier for resilient urban development (ASHRAE, 2025).

Current research methodologies for assessing urban health often rely on traditional simulations that fail to capture the fine-grained spatial-temporal dynamics of public sentiment (Bertrand et al., 2013). In New York City, health vulnerabilities were deeply intertwined with the spatial arrangement of critical infrastructure and socio-demographic factors (Kawlra & Sakamoto, 2023; Li & Huang, 2023). While periodic public polls offer some insight into regional satisfaction, they typically lack the geolocation granularity needed for targeted neighborhood-level interventions (Regional Plan Association, 2021). Furthermore, many existing studies emphasize socio-economic vulnerability without integrating the real-time human feedback identifying emerging building-level environmental concerns (Lamb et al., 2021). This methodological gap limits the ability of city administrators to implement targeted retrofit policies responding to the localized anxieties of New Yorkers regarding their indoor environments (Li & Huang, 2023).

Recent scholarship at the intersection of social sensing and performance-driven design suggests that user-generated text can be connected not only to public-health interpretation but also to city-scale building analytics, digital twins, adaptive urban governance, and environmental sustainability decision-making (Abbasabadi & Ashayeri, 2017; Abbasabadi & Ashayeri, 2024a, 2024b, 2024d, 2024e; Abbasabadi & Ashayeri, Eds., 2024). Studies linking social media to urban energy behavior and energy justice further show that such public data can reveal otherwise hidden patterns in building experience, vulnerability, and infrastructure performance (Abbasabadi & Ashayeri, 2024c, 2025; Ashayeri & Abbasabadi, 2024b).

This research objectives to address these critical gaps by quantifying the spatial-temporal shifts in public sentiment regarding indoor environmental quality in New York City. By applying fine-tuned Large Language Models to a comprehensive dataset of geolocated social media posts, the study seeks to map the evolution of public awareness from the onset of the pandemic to the current post-pandemic phase. The ultimate goal is to delineate a series of data-driven policy recommendations for building retrofits that align technical ventilation standards with human perceptions of safety. Through this systematic integration of social sensing and urban informatics, the research contributes to the development of more resilient, health-centric urban planning frameworks that prioritize the subjective well-being of the urban populace.

## **2- Materials and Methods**

We assembled a longitudinal corpus of 1,500 geotagged tweets posted between 1 January 2013 and 31 December 2023 that explicitly referenced indoor ventilation in New York City. Tweets

were identified through a keyword-based search strategy using ventilation-related terms such as *ventilation*, *airflow*, *air exchange*, and *air filtration*, and were retained only when geolocation metadata placed the post within the five boroughs of New York City. Retweets, exact duplicates, non-English posts, and tweets lacking interpretable textual content were excluded. Because geotagged Twitter data are known to reflect important demographic and behavioral biases, the resulting dataset was treated as a geographically bounded sample of public discourse rather than a representative measure of the broader population (Pavalanathan et al., 2015).

After collection, the tweets were preprocessed to prepare them for supervised text classification. URLs, usernames, and duplicated whitespace were normalized, while hashtags and punctuation were retained when they contributed semantic meaning. A manual relevance screening step was then applied to ensure that the final corpus captured explicit discussion of indoor ventilation rather than unrelated references to weather, outdoor conditions, or general building maintenance. The cleaned dataset was then organized for model development using stratified training, validation, and test splits so that class proportions were preserved across partitions.

For automated analysis, we fine-tuned RoBERTa, a robustly optimized transformer-based language model designed for improved performance on downstream natural language understanding tasks (Liu et al., 2019). RoBERTa was selected because its pretraining design improves on earlier BERT-style implementations and has demonstrated strong classification performance across short-text tasks. The model was trained to identify and classify ventilation-related discourse within the corpus, and performance was evaluated on the held-out test set using standard classification metrics, including precision, recall, and F1-score. Misclassified tweets were subsequently reviewed to identify recurrent sources of ambiguity, including sarcasm, abbreviated building references, and indirect mentions of indoor air conditions.

Methodologically, the present workflow is also consistent with recent NLP-based studies that decode indoor health perception and indoor air quality concerns from social media, as well as transformer-based analyses of equity and energy discourse in urban buildings (Ashayeri, 2024; Ashayeri, Piri, & Abbasabadi, 2024; Ashayeri & Abbasabadi, 2024b). The broader modeling logic also aligns with integrated urban energy and exposure frameworks and data-driven urban energy use modeling approaches that combine heterogeneous environmental and behavioral signals to improve urban inference (Abbasabadi et al., 2019; Abbasabadi & Ashayeri, 2019; Ashayeri & Abbasabadi, 2022; Ashayeri et al., 2021).

### **3- Results and Discussion**

As shown in Figure 1, ventilation-related tweet activity was very limited before 2018, then increased sharply from 2019 onward, with the strongest concentration in 2020. Negative sentiment rose from 33 tweets in 2019 to 92 in 2020, before declining to 68 in 2021, 54 in 2022, and 10 in 2023. Positive sentiment also increased substantially, peaking earlier at 71 tweets in 2019, then dropping to 26 in 2020 and stabilizing at lower but still elevated levels in 2021–2022. This pattern suggests that public discussion of indoor ventilation intensified during the COVID-19 period, when airborne transmission and the role of ventilation became central public-health concerns. The prominence of both positive and negative sentiment is consistent with the broader literature

showing that the pandemic reframed ventilation from a technical building issue into a visible public-health issue tied to risk, mitigation, and trust in indoor environments (Morawska et al., 2021; Morawska et al., 2024).

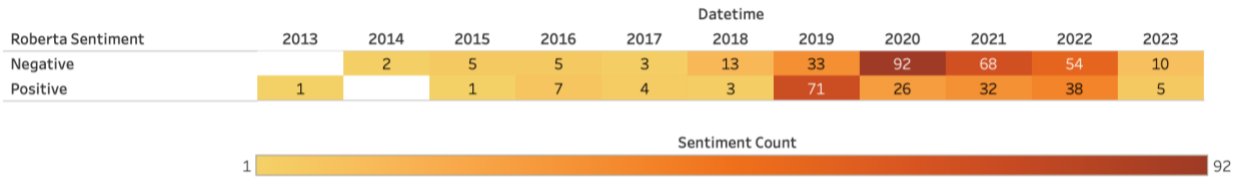


Figure 1. Annual distribution of RoBERTa-classified sentiment in ventilation-related tweets in New York City, 2013–2023.

The borough-level distribution in Figure 2 indicates that Manhattan dominated the discourse during the high-activity years, especially in 2019–2022. Negative tweets from Manhattan reached 50 in 2020, 51 in 2021, and 45 in 2022, while positive tweets were especially high in 2019 (67) before decreasing to 11 in 2020, then rebounding modestly in 2021 (23) and 2022 (27). Brooklyn showed the next most sustained activity, particularly for negative sentiment in 2020 (24), whereas Queens contributed more modest but consistent counts across both sentiment classes. The concentration in Manhattan likely reflects the borough’s dense mix of offices, transit nodes, commercial buildings, and public venues, where ventilation concerns would be more salient during and after the pandemic. At the same time, these spatial patterns should be interpreted carefully because geotagged Twitter data are not demographically or behaviorally representative of the wider population, and location-enabled users may systematically differ from other users (Pavalanathan & Eisenstein, 2015).

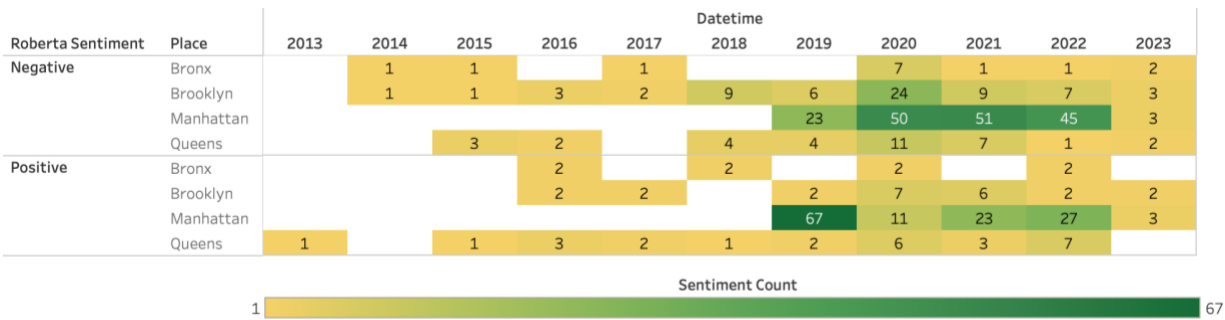


Figure 2. Spatial and temporal variation in RoBERTa-classified sentiment across New York City boroughs, 2013–2023.

Figure 3 shows that the RoBERTa model’s average confidence remained fairly stable across years, mostly clustering around 0.65–0.80, with a peak of 0.8651 for positive sentiment in 2017. During the core pandemic and post-pandemic years, confidence scores were notably consistent: for example, negative sentiment ranged from 0.6984 to 0.7598 between 2019 and 2023, while positive sentiment ranged from 0.6416 to 0.7103 over the same period. This stability suggests that the classifier performed reasonably consistently even as tweet volume changed substantially. However, years with very small counts, especially the early part of the series, should be treated more cautiously because isolated observations can inflate or distort yearly averages.

Average confidence per year across NYC boroughs



Figure 3. Average annual RoBERTa classification confidence for ventilation-related tweets in New York City boroughs, 2013–2023.

Taken together, the three figures suggest that public discussion of ventilation in New York City was not evenly distributed across time or space. Instead, it was concentrated in the pandemic era and especially in Manhattan, with negative sentiment persisting longer than positive sentiment after the 2020 peak. One plausible interpretation is that early positive discourse reflected optimism about ventilation as a protective intervention, whereas later negative discourse captured frustration over implementation gaps, uneven building conditions, or continued concern about indoor air risks. This interpretation aligns with post-pandemic literature arguing that ventilation became a lasting component of healthy-building discourse rather than a temporary emergency topic (Morawska et al., 2024). Even so, the findings are best read as an indicator of place-based online discourse rather than a direct measure of borough-wide public opinion. The Manhattan-centered clustering observed here is also compatible with research showing that social-media-derived signals can capture uneven urban burdens and place-specific dimensions of energy justice in New York City buildings (Ashayeri & Abbasabadi, 2024b).

#### 4. Future Research Directions

Future work should move beyond sentiment monitoring alone by integrating social sensing with performance-driven design and urban building energy modeling pipelines. Recent studies suggest strong opportunities to connect public discourse with digital twins, machine learning for urban building energy modeling, occupant-driven control strategies, and social-sensing-informed forecasting of urban energy behavior (Abbasabadi & Ashayeri, 2024b, 2024c, 2024d, 2024e, 2025). Such integration could help planners evaluate whether online concern about ventilation, exposure, or equity aligns with simulated building performance, neighborhood energy demand, or other measurable environmental indicators.

A second priority is to couple geolocated sentiment with richer spatial and design variables. Integrated energy-exposure models, PM2.5 prediction frameworks, microclimate-aware urban building energy modeling, and review work on simulation-to-reality gaps all indicate that future urban-health studies should link social-media signals to outdoor exposure, neighborhood morphology, and building-envelope dynamics (Ashayeri & Abbasabadi, 2022; Ashayeri et al., 2021; Worthy et al., 2025a, 2025b). In parallel, context-sensitive solar design and kinetic-facade optimization studies suggest that responsive building systems could eventually be evaluated not only for energy and daylight performance but also for their perceived health and comfort value in public discourse (Kamalisarvestani et al., 2023; Luitjohan et al., 2022). Extending the analysis to socioeconomic determinants of public health and residential energy use, and to community

resilience during health emergencies, would further strengthen the policy relevance of this line of research (Abbasabadi & Ashayeri, 2021; Ashayeri & Abbasabadi, 2021).

## 5. Conclusion

This study shows that ventilation-related discourse on Twitter in New York City was highly concentrated in the pandemic and immediate post-pandemic period, with the most substantial activity occurring between 2019 and 2022. The results indicate that Manhattan was the principal center of discussion and that negative sentiment remained more persistent than positive sentiment after the initial surge of attention. Together, these patterns suggest that indoor ventilation became an enduring urban public-health concern rather than a short-lived reaction to COVID-19, reflecting broader shifts in how indoor air quality is understood in public and policy discourse. At the same time, the findings should be interpreted with caution because geotagged X data capture a selective segment of users rather than the general population. Even with this limitation, the analysis demonstrates the value of combining geolocated social media with transformer-based sentiment classification to trace how health-related environmental concerns evolve across time and place. Future research could extend this work by integrating additional platforms, linking discourse patterns to specific policy or outbreak events, and comparing online sentiment with measured building or neighborhood-level ventilation conditions.

## Acknowledgement

X data were processed in accordance with the platform's Terms of Service and the academic research exemption. Tweet text and user handles were hashed with SHA-256 to prevent re-identification; only aggregate statistics are reported.

## References

- Abbasabadi, N., & Ashayeri, M. (2017). Towards an adaptive urbanism beyond hard control: The theories of Johnson and Lefebvre. In *ARCC architectural research addressing societal challenges* (Vol. 1, pp. 257-262). CRC Press/Taylor & Francis Group.
- Abbasabadi, N., & Ashayeri, M. (2019). Urban energy use modeling methods and tools: A review and an outlook. *Building and Environment*, 161, Article 106270. <https://doi.org/10.1016/j.buildenv.2019.106270>
- Abbasabadi, N., & Ashayeri, M. (2021). Socioeconomic determinants of public health and residential building energy use in Chicago. In *2021 International Conference: 27th World Congress of Architects (UIA 2021 RIO)* (Vol. II). ACSA.
- Abbasabadi, N., & Ashayeri, M. (Eds.). (2024a). *Artificial intelligence in performance-driven design: Theories, methods, and tools*. Wiley.

- Abbasabadi, N., & Ashayeri, M. (2024b). Digital twin for citywide energy modeling and management. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.
- Abbasabadi, N., & Ashayeri, M. (2024c). From Tweets to Energy Trends (TwEn): An exploratory framework for machine-learning-based forecasting of urban-scale energy behavior leveraging social media data. *Energy and Buildings*, 317, Article 114440. <https://doi.org/10.1016/j.enbuild.2024.114440>
- Abbasabadi, N., & Ashayeri, M. (2024d). Machine learning in urban building energy modeling. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.
- Abbasabadi, N., & Ashayeri, M. (2024e). Occupant-driven urban building energy efficiency via ambient intelligence. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.
- Abbasabadi, N., & Ashayeri, M. (2025). From Tweets to Energy Trends (TwEn2): Social-sensing-informed urban building energy modeling. *Frontiers in Energy Research*. <https://doi.org/10.3389/fenrg.2025.1688348>
- Abbasabadi, N., Ashayeri, M., Azari, R., Stephens, B., & Heidarinejad, M. (2019). An integrated data-driven framework for urban energy use modeling (UEUM). *Applied Energy*, 253, Article 113550. <https://doi.org/10.1016/j.apenergy.2019.113550>
- 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. <https://doi.org/10.1289/EHP.1510037>
- American Society of Heating, Refrigerating and Air-Conditioning Engineers. (2025). *ASHRAE position document on human health and wellness in the built environment*. <https://www.ashrae.org/file%20library/about/position%20documents/pd-on-human-health-and-wellness-in-the-built-environment-english.pdf>
- Ashayeri, M. (2024). Decoding global indoor health perception on social media via NLP and transformer deep learning. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.
- Ashayeri, M., & Abbasabadi, N. (2021). Energy justice, indoor air quality, and community health resilience against COVID-19 pandemic. In *Environments by design: Health, wellbeing and place; AMPS proceedings series* (Vol. 26.2). Architecture, Media, Politics, Society.
- Ashayeri, M., & Abbasabadi, N. (2022). A framework for integrated energy and exposure to ambient pollution assessment (iEnEx) toward low-carbon, healthy, and equitable cities. *Sustainable Cities and Society*, 78, Article 103647. <https://doi.org/10.1016/j.scs.2021.103647>
- Ashayeri, M., & Abbasabadi, N. (2024a). A hybrid physics-based machine learning approach for integrated energy and exposure modeling. In N. Abbasabadi & M. Ashayeri (Eds.),

- Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.
- Ashayeri, M., & Abbasabadi, N. (2024b). Unraveling energy justice in NYC urban buildings through social media sentiment analysis and transformer deep learning. *Energy and Buildings*, 306, Article 113914. <https://doi.org/10.1016/j.enbuild.2024.113914>
- Ashayeri, M., Abbasabadi, N., Heidarinejad, M., & Stephens, B. (2021). Predicting intraurban PM2.5 concentrations using enhanced machine learning approaches and incorporating human activity patterns. *Environmental Research*, 196, Article 110423. <https://doi.org/10.1016/j.envres.2020.110423>
- Ashayeri, M., Piri, S., & Abbasabadi, N. (2024). Exploring U.S. occupant perception toward indoor air quality via social media and NLP analysis. *Journal of Environmental Science and Public Health*, 8, 49-58. <https://doi.org/10.26502/jesph.96120205>
- Bertram, C., & Rehdanz, K. (2015). The role of urban green space for human well-being. *Ecological Economics*, 120, 139-152. <https://doi.org/10.1016/j.ecolecon.2015.10.013>
- Bertrand, K. Z., Bialik, M., Virdee, K., Gros, A., & Bar-Yam, Y. (2013). Sentiment in New York City: A high resolution spatial and temporal view. *arXiv*. <https://doi.org/10.48550/arXiv.1308.5010>
- California Department of Public Health. (2021, February 26). *Interim guidance for ventilation, filtration, and air quality in indoor environments*. <https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/COVID-19/Interim-Guidance-for-Ventilation-Filtration-and-Air-Quality-in-Indoor-Environments.aspx>
- Dengel, A. (2024, June 10). Why good indoor environmental quality can enhance wellbeing. *Safety Management Magazine*. British Safety Council. <https://www.britsafe.org/publications/safety-management-magazine/safety-management-magazine/2024/why-good-indoor-environmental-quality-can-enhance-wellbeing/>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171-4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Epifantsev, K., Bisht, S., Vanam, M., Prakash, A., & Sharma, M. (2024). Crowdsourced data for informed urban development: A social media crowdsourcing test. *BIO Web of Conferences*, 86, Article 01076. <https://doi.org/10.1051/bioconf/20248601076>
- Global Virus Network. (2023, July 3). *Improving ventilation and filtration for public health*. <https://gvn.org/improving-ventilation-and-filtration-for-public-health/>
- Hartmann, E. M. (2022, September 20). *Engineering and design factors for healthful built environments*. National Academy of Engineering. <https://www.nae.edu/281423/Engineering-and-Design-Factors-for-Healthful-Built-Environments>

- Hernandez-Martin, M., Del Ama Gonzalo, F., & Gonzalez-Lezcano, R. A. (2025). Indoor environmental quality to ensure the health and wellbeing of vulnerable people in residential buildings: A systematic review. *Frontiers in Built Environment*, 11, Article 1652527. <https://doi.org/10.3389/fbuil.2025.1652527>
- Ho, E., Schneider, M., Somanath, S., Yu, Y., & Thuvander, L. (2024). Sentiment and semantic analysis: Urban quality inference using machine learning algorithms. *iScience*, 27(7), Article 110192. <https://doi.org/10.1016/j.isci.2024.110192>
- Hong, L., Ye, G., Peng, Z., Wei, J., & Wu, C. (2022). Dynamic sentiment sensing of cities with social media data. *The Electronic Library*, 40(4), 413-434. <https://doi.org/10.1108/EL-03-2022-0064>
- Jiang, F., Ma, J., & Jin, Y. (2025). Unleashing the potential of large language models in urban data analytics: A review of emerging innovations and future research. *Smart Cities*, 8(6), Article 201. <https://doi.org/10.3390/smartcities8060201>
- Kamalisarvestani, S., Ashayeri, M., & Abbasabadi, N. (2023). An evolutionary multi-objective optimization tool for designing kinetic facades integrating daylight and lighting energy simulation. In *The Research-Design Interface: ARCC 2023 International Conference proceedings*. <http://www.arcc-arch.org/wp-content/uploads/2023/09/ARCC2023ProceedingsFINAL-PW.pdf>
- Kamel Boulos, M. N., Resch, B., Crowley, D. N., Breslin, J. G., Sohn, G., Burtner, R., Pike, W. A., Jezierski, E., & Chuang, K.-Y. S. (2011). Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: Trends, OGC standards and application examples. *International Journal of Health Geographics*, 10, Article 67. <https://doi.org/10.1186/1476-072X-10-67>
- Kawlra, G., & Sakamoto, K. (2023). Spatialising urban health vulnerability: An analysis of NYC's critical infrastructure during COVID-19. *Urban Studies*, 60(9), 1629-1649. <https://doi.org/10.1177/00420980211044304>
- Lamb, M. R., Kandula, S., & Shaman, J. (2021). Differential COVID-19 case positivity in New York City neighborhoods: Socioeconomic factors and mobility. *Influenza and Other Respiratory Viruses*, 15(2), 209-217. <https://doi.org/10.1111/irv.12816>
- Li, R., & Huang, Y. (2023). COVID-19 pandemic and minority health disparities in New York City: A spatial and temporal perspective. *Environment and Planning B: Urban Analytics and City Science*, 50(5), 1194-1211. <https://doi.org/10.1177/23998083221126525>
- Li, Z., Xia, L., Ren, X., Tang, J., Chen, T., Xu, Y., & Huang, C. (2025). Urban computing in the era of large language models. *ACM Transactions on Intelligent Systems and Technology*. <https://doi.org/10.1145/3768163>
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. *arXiv*. <https://openreview.net/forum?id=SyxS0T4tvS>
- Luitjohan, S., Ashayeri, M., & Abbasabadi, N. (2022). An optimization framework and tool for context-sensitive solar-driven design using cellular automata (SDCA). In *2022 Winter Simulation Conference (WSC)*. IEEE. <https://doi.org/10.23919/ANNSIM55834.2022.9859496>

- Morawska, L., Allen, J., Bahnfleth, W., Bluysen, P. M., Boerstra, A., Buonanno, G., Cao, J., Dancer, S. J., Floto, A., Franchimon, F., Greenhalgh, T., Haworth, C., Hogeling, J., Isaxon, C., Jimenez, J. L., Kurnitski, J., Li, Y., Loomans, M., Marks, G., Marr, L. C., Mazzarella, L., Melikov, A. K., Miller, S., Milton, D., Nazaroff, W., Nielsen, P. V., Noakes, C., Peccia, J., Prather, K., Querol, X., Sekhar, C., Seppanen, O., Tanabe, S., Tang, J. W., Tellier, R., Tham, K. W., Wargocki, P., & Yao, M. (2021). A paradigm shift to combat indoor respiratory infection. *Science*, 372(6543), 689-691. <https://doi.org/10.1126/science.abg2025>
- Morawska, L., Tang, J. W., Bahnfleth, W., Bluysen, P. M., Boerstra, A., Buonanno, G., Cao, J., Dancer, S., Floto, A., Franchimon, F., Greenhalgh, T., Haworth, C., Hogeling, J., Isaxon, C., Jimenez, J. L., Kurnitski, J., Li, Y., Loomans, M., Marks, G., Marr, L. C., Melikov, A. K., Miller, S., Nazaroff, W., Nielsen, P. V., Noakes, C., Peccia, J., Prather, K., Querol, X., Sekhar, C., Seppanen, O., Tanabe, S., Tellier, R., Tham, K. W., Wargocki, P., & Yao, M. (2024). Lessons from the COVID-19 pandemic for ventilation and indoor air quality. *Science*, 385(6707), 396-401. <https://doi.org/10.1126/science.adp2241>
- Morawska, L., Tang, J. W., Bahnfleth, W., Bluysen, P. M., Boerstra, A., Buonanno, G., Cao, J., Dancer, S., Floto, A., Franchimon, F., Haworth, C., Hogeling, J., Isaxon, C., Jimenez, J. L., Kurnitski, J., Li, Y., Loomans, M., Marks, G., Marr, L. C., Mazzarella, L., Melikov, A. K., Miller, S., Milton, D. K., Nazaroff, W., Nielsen, P. V., Noakes, C., Peccia, J., Prather, K., Querol, X., Sekhar, C., Seppanen, O., Tanabe, S.-I., Tellier, R., Tham, K. W., Wargocki, P., Wiedensohler, A., & Zhu, Y. (2020). How can airborne transmission of COVID-19 indoors be minimised? *Environment International*, 142, Article 105832. <https://doi.org/10.1016/j.envint.2020.105832>
- Pavalanathan, U., Eisenstein, J., & Lamba, H. (2015). Confounds and consequences in geotagged Twitter data. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 2138-2148). Association for Computational Linguistics. <https://aclanthology.org/D15-1256/>
- Pease, L. F., Wang, N., Salsbury, T. I., Underhill, R. M., Flaherty, J. E., Vlachokostas, A., Kulkarni, G., & James, D. P. (2021). Investigation of potential aerosol transmission and infectivity of SARS-CoV-2 through central ventilation systems. *Building and Environment*, 197, Article 107633. <https://doi.org/10.1016/j.buildenv.2021.107633>
- Perera, I., Hewage, K., Rana, A., & Sadiq, R. (2025). Combining energy performance and indoor environmental quality (IEQ) in buildings: A systematic review on common IEQ guidelines and energy codes in North America. *Energies*, 18(7), Article 1740. <https://doi.org/10.3390/en18071740>
- Rashid, M. T., & Wang, D. (2021). CovidSens: A vision on reliable social sensing for COVID-19. *Artificial Intelligence Review*, 54(1), 1-25. <https://doi.org/10.1007/s10462-020-09852-3>
- Regional Plan Association. (2021, March 25). *New York metro area issues survey I*. <https://rpa.org/work/reports/new-york-metro-area-issues-survey-i>
- Rousseau, S., & Deschacht, N. (2020). Public awareness of nature and the environment during the COVID-19 crisis. *Environmental and Resource Economics*, 76(4), 1149-1159. <https://doi.org/10.1007/s10640-020-00445-w>

- Samuelsson, K., Giusti, M., Peterson, G. D., Legeby, A., Brandt, S. A., & Barthel, S. (2018). Impact of environment on people's everyday experiences in Stockholm. *Landscape and Urban Planning*, 171, 7-17. <https://doi.org/10.1016/j.landurbplan.2017.11.009>
- Shi, D., Ye, X., Yang, J., Li, W., Xu, X., & Cai, Y. (2021). *Spatial-temporal data science for sustainable urban development*. Springer.
- Son, J. S., & Thill, J.-C. (2018). Is your city economic, cultural, or political? Recognition of city image based on multidimensional scaling of quantified web pages. In J.-C. Thill (Ed.), *Innovations in urban and regional systems* (pp. 63-95). Springer. [https://doi.org/10.1007/978-3-319-75998-2\\_4](https://doi.org/10.1007/978-3-319-75998-2_4)
- United States Environmental Protection Agency. (2026, March 18). *Preventing the spread of respiratory viruses in public indoor spaces*. <https://www.epa.gov/indoor-air-quality-iaq/preventing-spread-respiratory-viruses-public-indoor-spaces>
- Worthy, A., Ashayeri, M., Abbasabadi, N., & Julian, M. (2025a). Bridging the simulation-to-reality gap: A comprehensive review of microclimate integration in urban building energy modeling (UBEM). *Energy and Buildings*. <https://doi.org/10.1016/j.enbuild.2025.115392>
- Worthy, A., Ashayeri, M., & Abbasabadi, N. (2025b). Leveraging Earth observational data and machine learning to enhance urban building energy modeling with microclimate effects. *Sustainable Cities and Society*. <https://doi.org/10.1016/j.scs.2025.106544>
- Zhang, W., Deng, Y., Liu, B., Pan, S. J., & Bing, L. (2024). Sentiment analysis in the era of large language models: A reality check. In *Findings of the Association for Computational Linguistics: NAACL 2024* (pp. 3881-3906). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.findings-naacl.246>