

Analisis Big Data Untuk Pemantauan Kualitas Udara: Pendekatan, Implementasi, dan Tantangan dalam Studi Lingkungan

Muhamad Fuat Asnawi^{1*}, Nur Fitriyanto², M. Agoeng Pamoengkas³

¹Fakultas Teknik dan Ilmu Komputer, Program Studi Manajemen Informatika, Universitas Sains Al-Qur'an, Wonosobo, Indonesia

^{2,3}Fakultas Ilmu Komputer, Program Studi Informatika, Universitas Amikom Yogyakarta, Indonesia

Email: ¹fuatasnawi@unsiq.ac.id, ²nur.fitriyanto@students.amikom.ac.id, ³agoeng@students.amikom.ac.id

Email Penulis Korespondensi: fuatasnawi@unsiq.ac.id

Abstrak-Penelitian ini membahas penerapan big data dalam pemantauan kualitas udara, dengan fokus pada pendekatan, implementasi, dan tantangan yang dihadapi dalam pengelolaan kualitas udara di lingkungan perkotaan. Teknologi big data memungkinkan pengumpulan data dari berbagai sumber seperti sensor udara, stasiun pemantauan, dan data satelit, yang kemudian diolah menggunakan model machine learning untuk memprediksi kualitas udara secara lebih akurat. Studi ini juga mengidentifikasi tantangan utama dalam penerapan big data, termasuk masalah akurasi data, integrasi data dari berbagai sumber, serta kebutuhan infrastruktur komputasi yang lebih efisien. Dengan kemajuan teknologi seperti 5G dan edge computing, sistem pemantauan kualitas udara berbasis big data diprediksi akan semakin meningkat efisiensinya. Hasil dari kajian ini diharapkan dapat memberikan wawasan baru dalam upaya meningkatkan kualitas udara di kota-kota besar.

Kata Kunci: Big Data, Kualitas Udara, Machine Learning, Sensor IoT

Abstract-This study discusses the application of big data in air quality monitoring, focusing on approaches, implementation, and challenges encountered in managing air quality in urban environments. Big data technology enables the collection of data from various sources such as air sensors, monitoring stations, and satellite data, which are then processed using machine learning models to predict air quality more accurately. This study also identifies key challenges in implementing big data, including data accuracy issues, integration of data from multiple sources, and the need for more efficient computational infrastructure. With advancements in technologies such as 5G and edge computing, big data-based air quality monitoring systems are predicted to become increasingly efficient. The findings of this study are expected to provide new insights into efforts to improve air quality in major cities.

Keywords: Big Data, Air Quality, Machine Learning, IoT Sensors

1. PENDAHULUAN

Kualitas udara menjadi salah satu isu lingkungan paling mendesak di seluruh dunia. Menurut Organisasi Kesehatan Dunia (WHO), polusi udara menyebabkan lebih dari 4,2 juta kematian per tahun, dengan sebagian besar kematian terjadi di negara-negara berpenghasilan rendah dan menengah (Motlagh et al., 2020). Polusi udara juga berhubungan dengan berbagai masalah kesehatan akut dan kronis seperti penyakit pernapasan dan kardiovaskular (Motlagh et al., 2020). Meningkatnya urbanisasi, terutama di negara berkembang, memperparah masalah ini, dengan lebih dari 90% populasi dunia tinggal di daerah yang melebihi ambang batas kualitas udara yang aman (Ren et al., 2020). Dalam beberapa dekade terakhir, teknologi big data telah muncul sebagai alat penting dalam memantau dan menganalisis kualitas udara. Teknologi ini memungkinkan pengumpulan data dalam skala besar dari berbagai sumber seperti sensor udara, stasiun pemantauan, serta data satelit (Ren et al., 2020). Teknologi ini juga membuka peluang baru untuk membuat keputusan yang lebih akurat dan cepat dalam pengelolaan kualitas udara (Sun et al., 2021). Misalnya, pemantauan kualitas udara secara real-time dengan integrasi sensor dalam jaringan luas memungkinkan identifikasi pola polusi udara yang lebih rinci di tingkat lokal (Motlagh et al., 2020).

Penerapan big data dalam pemantauan kualitas udara membutuhkan pendekatan multidisiplin yang menggabungkan pengolahan data dalam jumlah besar, pemodelan prediktif, serta integrasi sistem pemantauan. Salah satu tantangan utama adalah bagaimana menyelaraskan data yang berasal dari berbagai sumber dan memiliki karakteristik berbeda-beda, seperti perbedaan dalam ketelitian pengukuran dan kualitas data (Hashem et al., 2023; Sun et al., 2021). Misalnya, sensor murah yang terpasang di transportasi umum atau infrastruktur perkotaan sering kali menghasilkan data yang kurang akurat tanpa kalibrasi yang tepat dengan stasiun pemantauan profesional (Motlagh et al., 2020). Selain itu, kualitas data sangat dipengaruhi oleh faktor lingkungan seperti kelembaban dan suhu, yang dapat menyebabkan variasi hasil pengukuran. Untuk mengatasi masalah ini, berbagai pendekatan seperti kalibrasi kolaboratif dan pengumpulan data yang oportunistik perlu diterapkan (Motlagh et al., 2020). Misalnya, sensor yang terpasang pada kendaraan atau alat transportasi umum dapat dikalibrasi secara otomatis saat melewati stasiun pengukuran profesional, sehingga meningkatkan akurasi hasil pemantauan (Motlagh et al., 2020).

Secara keseluruhan, penelitian ini bertujuan untuk mengeksplorasi lebih lanjut bagaimana pendekatan big data dapat diterapkan dalam memantau dan meningkatkan kualitas udara, serta tantangan yang dihadapi dalam penerapannya. Dengan menggabungkan berbagai metode analisis data dan model prediktif, diharapkan solusi yang lebih efektif dapat dikembangkan untuk mengatasi masalah polusi udara di lingkungan perkotaan (Chen et al., 2023; D. Liu & Shi, 2022).

2. KERANGKA TEORI

Skala implementasi teknologi big data dalam pemantauan kualitas udara menghadapi banyak kendala, termasuk kebutuhan akan infrastruktur komputasi yang efisien, pemrosesan data real-time, dan integrasi data dari sumber-sumber

yang berbeda. Perkembangan teknologi seperti 5G dan komputasi edge diharapkan dapat mengatasi beberapa dari tantangan ini dengan memungkinkan pengolahan data dalam skala besar secara lebih cepat dan efisien (Motlagh et al., 2020; Ren et al., 2020). Selain itu, masalah keamanan data dan privasi juga menjadi perhatian penting, terutama ketika data dikumpulkan dalam jumlah besar dari area publik (Hashem et al., 2023; J. Liu et al., 2020). Dalam konteks perkotaan, data kualitas udara yang dihasilkan oleh sistem pemantauan berbasis big data dapat digunakan untuk mendukung kebijakan lingkungan, seperti peringatan dini terhadap peristiwa polusi udara yang berbahaya dan penetapan rute transportasi yang lebih bersih (Motlagh et al., 2020). Selain itu, data ini juga memungkinkan pemerintah untuk melakukan tindakan mitigasi yang lebih cepat dan efisien dalam menghadapi peningkatan konsentrasi polutan di daerah perkotaan (Ren et al., 2020). Dengan demikian, pendekatan berbasis big data tidak hanya meningkatkan pemahaman terhadap pola polusi udara, tetapi juga mempercepat respons dalam pengambilan keputusan kebijakan (Cao et al., 2020; Sun et al., 2021).

3. METODE PENELITIAN

Penelitian ini bertujuan untuk mengidentifikasi dan menganalisis pendekatan berbasis big data yang digunakan dalam pemantauan kualitas udara, jenis data yang digunakan dalam sistem ini, serta tantangan dan evaluasi performa implementasi teknologi tersebut. Literature review ini akan meninjau 35 artikel dari berbagai sumber yang relevan dan diterbitkan dalam rentang waktu 2020 hingga 2024.

Strategi pencarian dilakukan dengan menggunakan beberapa kata kunci utama yaitu "air quality", "big data", dan variasi terkait lainnya. Pencarian ini difokuskan pada artikel ilmiah, prosiding konferensi, serta laporan penelitian yang diterbitkan dalam rentang waktu 2020 hingga 2024. Basis data yang digunakan mencakup platform akademik utama seperti IEEE Xplore, Scopus, ScienceDirect, Web of Science, MDPI dan lain sebagainya. Proses pencarian dilakukan dengan memadukan kata kunci utama dengan variasi istilah terkait, seperti "air quality monitoring using big data", "predictive analytics for air quality", "big data air pollution control", dan "sensor technology in air quality monitoring". Kriteria inklusi meliputi artikel yang fokus pada penerapan big data dalam pemantauan kualitas udara, studi kasus di wilayah perkotaan, serta artikel yang meninjau teknologi pemrosesan data besar dalam konteks lingkungan. Artikel yang tidak relevan atau diterbitkan sebelum tahun 2020 dikecualikan dari analisis ini. Setelah tahap pencarian awal, hasil yang didapat disaring lebih lanjut berdasarkan abstrak dan kesesuaian topik untuk memastikan artikel yang digunakan memberikan kontribusi signifikan terhadap kajian literatur tentang kualitas udara dan teknologi big data.

Berikut ini adalah RQ dari literature review yang kami analisa:

- 1) RQ1: Bagaimana teknologi big data dapat digunakan untuk meningkatkan akurasi prediksi dan pemantauan kualitas udara secara real-time di berbagai lingkungan perkotaan?
- 2) RQ2: Apa saja jenis data yang paling sering digunakan dalam pendekatan berbasis big data untuk pemantauan kualitas udara, dan bagaimana data tersebut diintegrasikan dari berbagai sumber seperti sensor, data satelit, dan data publik?
- 3) RQ3: Apa tantangan utama yang dihadapi dalam penerapan metode big data untuk pemantauan kualitas udara, dan bagaimana tantangan tersebut dapat diatasi untuk meningkatkan efisiensi dan keandalan sistem?

4. HASIL

Teknologi big data menawarkan peluang besar dalam meningkatkan akurasi prediksi kualitas udara, terutama dengan memanfaatkan data yang bersifat real-time dari berbagai sensor yang tersebar di area perkotaan. Misalnya, studi dari Motlagh menunjukkan bahwa penerapan jaringan sensor yang luas, baik pada infrastruktur kota maupun kendaraan umum, memungkinkan pengumpulan data kualitas udara yang lebih detail dan bervariasi (Motlagh et al., 2020). Studi ini juga menunjukkan bagaimana pemrosesan data secara kolaboratif melalui jaringan 5G dan edge computing dapat mendukung pemantauan kualitas udara yang lebih cepat dan akurat di kota-kota besar (Jo et al., 2023; J. Liu et al., 2020). Selain itu, data satelit juga dimanfaatkan untuk melengkapi data dari darat, seperti yang dibahas dalam studi oleh Ren, di mana integrasi data satelit memungkinkan analisis yang lebih holistik tentang distribusi polutan di atmosfer (Ren et al., 2020).

Studi lain, seperti yang diungkapkan oleh Sun, membahas bahwa pemodelan berbasis machine learning memainkan peran kunci dalam memprediksi konsentrasi polutan seperti PM_{2.5} dan NO₂ (Sun et al., 2021). Algoritma seperti long short-term memory (LSTM) telah banyak digunakan dalam memprediksi tren kualitas udara berdasarkan data historis dan real-time. Model AI ini memungkinkan prediksi lebih tepat dalam jangka pendek maupun jangka Panjang (K. Bai et al., 2022; Koo et al., 2023). Penelitian lain juga menunjukkan bahwa pemodelan prediktif berbasis big data mampu mendeteksi perubahan pola polusi dengan tingkat akurasi yang lebih tinggi dibandingkan metode konvensional (Chen et al., 2023). Namun, efektivitas teknologi big data dalam prediksi kualitas udara tidak terlepas dari tantangan seperti skala implementasi yang besar dan kebutuhan akan pemrosesan data yang cepat. Liu menyebutkan bahwa pemantauan udara pada skala besar memerlukan sensor berkualitas tinggi serta mekanisme pemrosesan data yang mampu menangani data besar secara real-time (D. Liu & Shi, 2022). Beberapa studi menyoroti pentingnya pengembangan platform kolaboratif yang dapat menyatukan data dari berbagai sumber untuk menghasilkan gambaran yang lebih akurat tentang kualitas udara di kota-kota besar (De et al., 2021; Kaginalkar et al., 2022; D. Liu & Shi, 2022; Ren et al., 2020). Dengan kemajuan teknologi seperti 5G dan multi-access edge computing, sistem pemantauan kualitas udara berbasis big data diprediksi akan semakin meningkat efisiensinya (Kaginalkar et al., 2021; Liang et al., 2023; D. Liu & Shi, 2022).

Jenis data yang paling umum digunakan dalam pendekatan berbasis big data untuk pemantauan kualitas udara berasal dari berbagai sumber, termasuk data sensor, data satelit, serta data meteorologi dan publik. Studi oleh Ren menunjukkan bahwa data satelit memainkan peran penting dalam memetakan distribusi polutan di atmosfer (Ren et al., 2020). Selain itu, sensor udara yang terpasang pada kendaraan atau tempat umum memungkinkan pengumpulan data secara real-time, dengan cakupan area yang lebih luas, seperti yang dibahas dalam penelitian oleh Sun (Sun et al., 2021). Selain itu, data meteorologi seperti suhu, kelembaban, dan kecepatan angin digunakan untuk mengkontekstualisasikan konsentrasi polutan udara (D. Liu & Shi, 2022).

Data yang dihasilkan dari sensor udara memiliki beragam karakteristik tergantung pada jenis sensor yang digunakan. Motlagh mengungkapkan bahwa sensor biaya rendah dapat diintegrasikan ke dalam jaringan kota untuk menangkap data kualitas udara secara kontinu, namun akurasi data ini bergantung pada proses kalibrasi yang rutin (Ma et al., 2020; Motlagh et al., 2020; Wu et al., 2020). Data dari sensor-sensor ini kemudian diolah melalui sistem komputasi edge dan cloud, memungkinkan pengambilan keputusan yang lebih cepat untuk berbagai aplikasi, termasuk peringatan dini polusi udara (Bachechi et al., 2022; Borrajo & Cao, 2020; Chen et al., 2023; Motlagh et al., 2020). Penggunaan sensor IoT juga disebutkan oleh Liu sebagai bagian penting dari infrastruktur smart city untuk pengumpulan data kualitas udara yang lebih presisi (Kok et al., 2020; D. Liu & Shi, 2022; Tshering et al., 2021).

Proses integrasi data dari berbagai sumber ini menjadi tantangan tersendiri. Dalam penelitian yang dibahas oleh Hashem, pentingnya pengolahan data dari berbagai sumber dengan teknik data fusion menjadi fokus utama (Hashem et al., 2023). Menggabungkan data dari sensor darat dan data satelit membutuhkan metode pemrosesan yang mampu menyelaraskan perbedaan format, waktu, dan resolusi data (K. Bai et al., 2022; Ren et al., 2020; Sun et al., 2021). Penggunaan teknik seperti spatiotemporal analysis juga membantu dalam mengintegrasikan berbagai jenis data untuk menghasilkan pemahaman yang lebih menyeluruh mengenai kualitas udara di suatu wilayah (Chen et al., 2023).

Salah satu tantangan utama dalam penerapan metode big data untuk pemantauan kualitas udara adalah masalah kualitas dan akurasi data, terutama ketika menggunakan sensor berbiaya rendah. Seperti yang diungkapkan oleh Motlagh, sensor murah sering kali rentan terhadap pengaruh kondisi lingkungan seperti suhu dan kelembaban, sehingga data yang dihasilkan bisa kurang akurat tanpa kalibrasi yang tepat (Iskandaryan et al., 2020; Liang et al., 2023; Motlagh et al., 2020). Studi lain juga menunjukkan bahwa kesenjangan dalam keakuratan sensor ini menjadi kendala dalam implementasi jaringan pemantauan kualitas udara yang lebih luas (Ren et al., 2020). Salah satu solusi yang diusulkan adalah penerapan kalibrasi kolaboratif, di mana sensor-sensor ini secara otomatis dikalibrasi ketika berada dalam jarak dekat dengan stasiun pemantauan profesional (Motlagh et al., 2020).

Selain tantangan dalam kualitas data, masalah integrasi data dari berbagai sumber juga menjadi hambatan signifikan. Data dari sensor, satelit, dan publik sering kali memiliki format yang berbeda dan tidak seragam, yang membuat proses integrasi menjadi kompleks (Sun et al., 2021). Dalam studi yang dilakukan oleh Ren, proses data fusion diusulkan sebagai solusi untuk mengatasi perbedaan ini, namun membutuhkan infrastruktur teknologi yang lebih maju untuk memproses volume data yang besar secara real-time (X. Bai et al., 2024; Chen et al., 2023; Fu et al., 2023; Himeur et al., 2023; Nam et al., 2022; Ren et al., 2020). Tantangan lainnya adalah keterbatasan dalam infrastruktur komputasi dan jaringan, terutama di wilayah dengan sumber daya yang terbatas.

Untuk mengatasi tantangan-tantangan ini, beberapa studi menyarankan penggunaan teknologi 5G dan edge computing untuk mempercepat proses pemrosesan data (Kaginalkar et al., 2022; Letaief et al., 2022; J. Liu et al., 2020; Ma et al., 2020; Sokhi et al., 2022; Zhou et al., 2023). Dengan teknologi ini, data dapat diproses secara lokal tanpa perlu dikirim ke pusat data yang jauh, sehingga memungkinkan respons yang lebih cepat terhadap peristiwa polusi udara (Barthwal et al., 2023; Chang et al., 2023; Jufriansah et al., 2023; Mohammadi et al., 2024; Ren et al., 2020). Selain itu, penerapan sistem berbasis AI dan machine learning memungkinkan pengolahan data yang lebih efisien serta deteksi polusi udara yang lebih akurat, sebagaimana dibahas oleh sun (Asgari et al., 2022; Chen et al., 2023; Muthukumar et al., 2022; Sun et al., 2021).

5. KESIMPULAN

Dalam penerapan big data untuk pemantauan kualitas udara, teknologi ini telah menunjukkan potensi besar dalam meningkatkan akurasi prediksi dan analisis kualitas udara, terutama di area perkotaan. Dengan menggabungkan data dari berbagai sumber seperti sensor IoT, data satelit, dan data meteorologi, teknologi big data memungkinkan pengolahan informasi dalam skala besar secara lebih cepat dan efisien. Namun, tantangan seperti kualitas data dari sensor murah, keterbatasan infrastruktur komputasi, dan kesulitan dalam integrasi data masih menjadi hambatan utama. Solusi yang diusulkan meliputi penggunaan teknologi 5G, komputasi edge, serta teknik data fusion untuk menyelaraskan perbedaan format dan waktu dalam data. Dengan solusi ini, diharapkan sistem pemantauan kualitas udara dapat memberikan hasil yang lebih andal dan cepat untuk mendukung kebijakan lingkungan dan mitigasi polusi udara.

DAFTAR PUSTAKA

- Asgari, M., Yang, W., & Farnaghi, M. (2022). Spatiotemporal data partitioning for distributed random forest algorithm: Air quality prediction using imbalanced big spatiotemporal data on spark distributed framework. *Environmental Technology and Innovation*, 27. <https://doi.org/10.1016/j.eti.2022.102776>
- Bachechi, C., Po, L., & Rollo, F. (2022). Big Data Analytics and Visualization in Traffic Monitoring. *Big Data Research*, 27. <https://doi.org/10.1016/j.bdr.2021.100292>

- Bai, K., Li, K., Ma, M., Li, K., Li, Z., Guo, J., Chang, N. Bin, Tan, Z., & Han, D. (2022). LGHAP: The Long-Term Gap-free High-resolution Air Pollutant concentration dataset, derived via tensor-flow-based multimodal data fusion. *Earth System Science Data*, 14(2), 907–927. <https://doi.org/10.5194/essd-14-907-2022>
- Bai, X., Li, P., Zhou, W., Wu, H., Li, C., & Zhou, Z. (2024). Temporal Refinement of Major Primary Air Pollutant Emissions Based on Electric Power Big Data: A Case of the Cement Industry in Tangshan City. *Atmosphere*, 15(8). <https://doi.org/10.3390/atmos15080895>
- Barthwal, A., Acharya, D., & Lohani, D. (2023). Prediction and analysis of particulate matter (PM2.5 and PM10) concentrations using machine learning techniques. *Journal of Ambient Intelligence and Humanized Computing*, 14(3), 1323–1338. <https://doi.org/10.1007/s12652-021-03051-w>
- Borrajó, L., & Cao, R. (2020). Big-but-biased data analytics for air quality. *Electronics (Switzerland)*, 9(9), 1–11. <https://doi.org/10.3390/electronics9091551>
- Cao, Z., Wu, Z., Li, S., Ma, W., Deng, Y., Sun, H., & Guan, W. (2020). Exploring spatiotemporal variation characteristics of exceedance air pollution risk using social media big data. *Environmental Research Letters*, 15(11). <https://doi.org/10.1088/1748-9326/abb62>
- Chang, Y. S., Abimananan, S., Shukla, S., & Satheesh, D. (2023). Forecasting Continuous Air Pollution with Cloud-Based Big Data Platform using Spatial-Temporal Visualization. 2023 IEEE 8th International Conference for Convergence in Technology, I2CT 2023. <https://doi.org/10.1109/I2CT57861.2023.10126153>
- Chen, J. T., Hu, C. F., & Wu, C. C. (2023). Teaching Big Data Concepts through Air Pollution Data: A High School Instructional Module. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 2, 651. <https://doi.org/10.1145/3587103.3594194>
- De, S., Jassat, U., Wang, W., Perera, C., & Moessner, K. (2021). Inferring Latent Patterns in Air Quality from Urban Big Data. *IEEE Internet of Things Magazine*, 4(1), 20–27. <https://doi.org/10.1109/IOTM.0011.2000071>
- Fu, L., Li, J., & Chen, Y. (2023). An innovative decision making method for air quality monitoring based on big data-assisted artificial intelligence technique. *Journal of Innovation and Knowledge*, 8(2). <https://doi.org/10.1016/j.jik.2022.100294>
- Hashem, I. A. T., Usmani, R. S. A., Almutairi, M. S., Ibrahim, A. O., Zakari, A., Alotaibi, F., Alhashmi, S. M., & Chiroma, H. (2023). Urban Computing for Sustainable Smart Cities: Recent Advances, Taxonomy, and Open Research Challenges. In *Sustainability (Switzerland)* (Vol. 15, Issue 5). MDPI. <https://doi.org/10.3390/su15053916>
- Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., Bensaali, F., & Amira, A. (2023). AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. *Artificial Intelligence Review*, 56(6), 4929–5021. <https://doi.org/10.1007/s10462-022-10286-2>
- Iskandaryan, D., Ramos, F., & Trilles, S. (2020). Air quality prediction in smart cities using machine learning technologies based on sensor data: A review. In *Applied Sciences (Switzerland)* (Vol. 10, Issue 7). MDPI AG. <https://doi.org/10.3390/app10072401>
- Jo, J., Kim, G., Kwak, J., Jeong, I., Ku, D., & Lee, S. (2023). Air Quality Modeling Using Real-time Urban Big Data. *Chemical Engineering Transactions*, 106, 229–234. <https://doi.org/10.3303/CET23106039>
- Jufriansah, A., Khusnani, A., Pramudya, Y., Sya'bania, N., Leto, K. T., Hikmatiar, H., & Saputra, S. (2023). AI Big Data System to Predict Air Quality for Environmental Toxicology Monitoring. *Journal of Novel Engineering Science and Technology*, 2(01), 21–25. <https://doi.org/10.56741/jnest.v2i01.314>
- Kaginalkar, A., Kumar, S., Gargava, P., Kharkar, N., & Niyogi, D. (2022). SmartAirQ: A Big Data Governance Framework for Urban Air Quality Management in Smart Cities. In *Frontiers in Environmental Science* (Vol. 10). Frontiers Media S.A. <https://doi.org/10.3389/fenvs.2022.785129>
- Kaginalkar, A., Kumar, S., Gargava, P., & Niyogi, D. (2021). Changing Paradigm of Urban Air Quality Management: Opportunities and Challenges of Frontier 2 Technologies in the context of Smart Cities 3.
- Kok, C. H., Azlan, M. A. M., & Ong, S. E. (2020). IoT based Low Cost Distributed Air Quality Monitoring System for Big Data Collection. *Proceedings - 10th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2020*, 7–10. <https://doi.org/10.1109/ICCSCE50387.2020.9204954>
- Koo, Y. S., Choi, Y., & Ho, C. -H. (2023). Air Quality Forecasting Using Big Data and Machine Learning Algorithms. In *Asia-Pacific Journal of Atmospheric Sciences* (Vol. 59, Issue 5, pp. 529–530). Korean Meteorological Society. <https://doi.org/10.1007/s13143-023-00347-z>
- Letaief, K. B., Shi, Y., Lu, J., & Lu, J. (2022). Edge Artificial Intelligence for 6G: Vision, Enabling Technologies, and Applications. *IEEE Journal on Selected Areas in Communications*, 40(1), 5–36. <https://doi.org/10.1109/JSAC.2021.3126076>
- Liang, L., Daniels, J., Bailey, C., Hu, L., Phillips, R., & South, J. (2023). Integrating low-cost sensor monitoring, satellite mapping, and geospatial artificial intelligence for intra-urban air pollution predictions. *Environmental Pollution*, 331. <https://doi.org/10.1016/j.envpol.2023.121832>
- Liu, D., & Shi, Y. (2022). The Influence Mechanism of Urban Spatial Structure on Urban Vitality Based on Geographic Big Data: A Case Study in Downtown Shanghai. *Buildings*, 12(5). <https://doi.org/10.3390/buildings12050569>
- Liu, J., Li, T., Xie, P., Du, S., Teng, F., & Yang, X. (2020). Urban big data fusion based on deep learning: An overview. *Information Fusion*, 53, 123–133. <https://doi.org/10.1016/j.inffus.2019.06.016>
- Ma, J., Ding, Y., Cheng, J. C. P., Jiang, F., Tan, Y., Gan, V. J. L., & Wan, Z. (2020). Identification of high impact factors of air quality on a national scale using big data and machine learning techniques. *Journal of Cleaner Production*, 244. <https://doi.org/10.1016/j.jclepro.2019.118955>
- Mohammadi, F., Teiri, H., Hajizadeh, Y., Abdollahnejad, A., & Ebrahimi, A. (2024). Prediction of atmospheric PM2.5 level by machine learning techniques in Isfahan, Iran. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-52617-z>
- Motlagh, N. H., Petäjä, T., Kulmala, M., Trachoma, S., Lagerspetz, E., Nurmi, P., Li, X., Varjonen, S., Mineraud, J., Siekkinen, M., Rebeiro-Hargrave, A., & Hussein, T. (2020). Toward Massive Scale Air Quality Monitoring. *IEEE Communications Magazine*, 58(2), 54–59. <https://doi.org/10.1109/MCOM.001.1900515>
- Muthukumar, P., Cocom, E., Nagrecha, K., Comer, D., Burga, I., Taub, J., Calvert, C. F., Holm, J., & Pourhomayoun, M. (2022). Predicting PM2.5 atmospheric air pollution using deep learning with meteorological data and ground-based observations and remote-sensing satellite big data. *Air Quality, Atmosphere and Health*, 15(7), 1221–1234. <https://doi.org/10.1007/s11869-021-01126-3>



- Nam, S., Shin, M. Y., Han, J. Y., Moon, S. Y., Kim, J. Y., Tchah, H., & Lee, H. (2022). Correlation between air pollution and prevalence of conjunctivitis in South Korea using analysis of public big data. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-13344-5>
- Ren, X., Zhao, Y., Wu, L., Jiang, J., Zhang, C., Tang, X., Han, L., Zhu, M., & Wang, Z. (2020). Towards efficient digital governance of city air pollution using technique of big atmospheric environmental data. *IOP Conference Series: Earth and Environmental Science*, 502(1). <https://doi.org/10.1088/1755-1315/502/1/012031>
- Sokhi, R. S., Moussiopoulos, N., Baklanov, A., Bartzis, J., Coll, I., Finardi, S., Friedrich, R., Geels, C., Grönholm, T., Halenka, T., Ketzel, M., Maragkidou, A., Matthias, V., Moldanova, J., Ntziachristos, L., Schäfer, K., Suppan, P., Tsegas, G., Carmichael, G., ... Kukkonen, J. (2022). Advances in air quality research - current and emerging challenges. In *Atmospheric Chemistry and Physics* (Vol. 22, Issue 7, pp. 4615–4703). Copernicus GmbH. <https://doi.org/10.5194/acp-22-4615-2022>
- Sun, W., Hou, Y., & Guo, L. (2021). Big data revealed relationship between air pollution and manufacturing industry in China. *Natural Hazards*, 107(3), 2533–2553. <https://doi.org/10.1007/s11069-020-04495-7>
- Tshering, Y., Tamrakar, S. R., Gontia, S., & Baral, S. (2021). IoT-based Platform with Big Data using Apache Kylin: Air Quality Monitoring System (AQMS). *World Journal of Research and Review*, 13(3). <https://doi.org/10.31871/wjrr.13.3.12>
- Wu, L., Chang, M., Wang, X., Hang, J., Zhang, J., Wu, L., & Shao, M. (2020). Development of the Real-time On-road Emission (ROE v1.0) model for street-scale air quality modeling based on dynamic traffic big data. *Geoscientific Model Development*, 13(1), 23–40. <https://doi.org/10.5194/gmd-13-23-2020>
- Zhou, P., Ma, J., Li, X., Zhao, Y., Yu, K., Su, R., Zhou, R., Wang, H., & Wang, G. (2023). The long-term and short-term effects of ambient air pollutants on sleep characteristics in the Chinese population: big data analysis from real world by sleep records of consumer wearable devices. *BMC Medicine*, 21(1). <https://doi.org/10.1186/s12916-023-02801-1>