Applications of artificial intelligence in the field of air pollution: A bibliometric analysis

Qiangqiang Guo[1], Mengjuan Ren[1], Shouyuan Wu[1], Yajia Sun[1], Jianjian Wang[1], Qi Wang[2,3], Yanfang Ma[4], Xuping Song[1,5,6,7]* and Yaolong Chen[1,5,6,7]*

[1]School of Public Health, Lanzhou University, Lanzhou, China, [2]Department of Health Research Methods, Evidence and Impact, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada, [3]McMaster Health Forum, McMaster University, Hamilton, ON, Canada, [4]School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China, [5]Research Unit of Evidence-Based Evaluation and Guidelines, Chinese Academy of Medical Sciences (2021RU017), School of Basic Medical Sciences, Lanzhou University, Lanzhou, China, [6]Lanzhou University Institute of Health Data Science, Lanzhou, China, [7]World Health Organization Collaborating Center for Guideline Implementation and Knowledge Translation, Lanzhou, China

Background: Artificial intelligence (AI) has become widely used in a variety of fields, including disease prediction, environmental monitoring, and pollutant prediction. In recent years, there has also been an increase in the volume of research into the application of AI to air pollution. This study aims to explore the latest trends in the application of AI in the field of air pollution.

Methods: All literature on the application of AI to air pollution was searched from the Web of Science database. CiteSpace 5.8.R1 was used to analyze countries/regions, institutions, authors, keywords and references cited, and to reveal hot spots and frontiers of AI in atmospheric pollution.

Results: Beginning in 1994, publications on AI in air pollution have increased in number, with a surge in research since 2017. The leading country and institution were China ($N = 524$) and the Chinese Academy of Sciences ($N = 58$), followed by the United States ($N = 455$) and Tsinghua University ($N = 33$), respectively. In addition, the United States (0.24) and the England (0.27) showed a high degree of centrality. Most of the identified articles were published in journals related to environmental science; the most cited journal was *Atmospheric Environment*, which reached nearly 1,000 citations. There were few collaborations among authors, institutions and countries. The hot topics were machine learning, air pollution and deep learning. The majority of the researchers concentrated on air pollutant concentration prediction, particularly the combined use of AI and environmental science methods, low-cost air quality sensors, indoor air quality, and thermal comfort.

Conclusion: Researches in the field of AI and air pollution are expanding rapidly in recent years. The majority of scholars are from China and the United States, and the Chinese Academy of Sciences is the dominant research institution. The United States and the England contribute greatly to the development of the cooperation network. Cooperation among research institutions appears to be suboptimal, and strengthening cooperation could greatly benefit this field of research. The prediction of air pollutant

concentrations, particularly $PM_{2.5}$, low-cost air quality sensors, and thermal comfort are the current research hotspot.

Introduction

Air pollution is defined as all destructive effects from any sources which contribute to the pollution of the atmosphere and/or deterioration of the environment (1). It was initially regarded as a threat to respiratory health, however as air pollution research advanced, public health concerns were broadened to include many diseases (2–4), which caused a major threat to human public health. According to the World Health Organization's *COP26 Special Report on Climate Change and Health*, over 90% of people breathe unhealthy levels of outdoor air pollution, which is largely caused by the burning of fossil fuels, which is also driving climate change. It is the first time that air pollution has been listed as a cause of death in the report. Therefore, the World Health Organization launched the "let us breathe clean air" initiative in 2021 (5).

Artificial intelligence (AI) is a new technological science that studies and develops theories, methods, technologies and application systems for simulating, extending and expanding human intelligence (6). One of its main goals is to enable machines to perform complex tasks that would normally require human intelligence. Advantages of AI, such as automatic data extraction, efficiencies in terms of time and labor expenditures, convenience, and long-term sustainability (7, 8), allow for a wide range of applications, including disease prediction, weather forecasting, expert system construction, environmental monitoring, and pollutant prediction (9–14). In recent years, research on the application of AI in air pollution has increased (15). Hu et al. developed a random forest model, incorporating aerosol optical depth data, meteorological fields, and land use variables to estimate daily 24 h averaged ground-level fine particulate matter ($PM_{2.5}$) concentrations over the United States (16). Li et al. developed a novel long short-term memory neural network extended model to predict the air pollutant concentration (17). Huang and Kuo also developed a deep neural network model to apply the $PM_{2.5}$ forecasting system (18).

Bibliometrics is a discipline that uses mathematics, statistics, and other measurement methods to study the distribution structure, quantitative relationships, change patterns, and quantitative management of literature and intelligence. It then investigates specific structures, characteristics, and laws of science and technology, using the literature system and bibliometric characteristics as the object of study. Researchers have begun to use computers for bibliometric work as computers have become more popular. Modern methods and tools, such as computers, are used for data processing and analysis by establishing a systematic and standardized system of data sources and access to primary data. For example, Citation Space (known as "CiteSpace") is a Java-based information visualization software developed in 2004 by Professor Chaomei Chen of Drexel University's School of Information Science and Technology (USA). It allows the visualization of the structure, patterns and distribution of scientific knowledge. CiteSpace software not only analyzes co-citations in the literature, thus mining the citation space for knowledge clustering and distribution, but also provides the ability to analyze co-occurrence between other knowledge units, allowing for a better presentation and understanding of the progression of scientific knowledge (19). Article co-citation is said to constitute the co-citation relationship when two (or more) articles are simultaneously cited by one or more subsequent articles. When compared to other clusters, an object's silhouette value indicates how similar to its own cluster it is. Modularity and silhouette have values that range from −1 to 1, with values closer to 1 being deemed desirable. An indicator of a node's relevance in a network is called centrality. In their collaborative networks, any research with centrality ratings higher than 0.1 is regarded as influential. Citation bursts, which reveal brief periods of intense scholarly activity, are depicted as red rings. Researchers can identify important hotspots in a field based on these findings.

This study aims to explore the current status of the application of AI in the field of air pollution research using bibliometrics and visual analysis of CiteSpace software (20–22), including the distribution of countries or regions where the research was conducted, authors, and journals, as well as the research hotspots and trends. These analyses then lay the foundation for subsequent relevant research.

Materials and methods

Data collection

The data were obtained from the Web of Science (WOS) Core Collection database. Two researchers conducted the search independently on October 13, 2021. The search period was from inception of the database (1980) to October 12, 2021. The main search terms were "air pollution", "air contamination", "atmosphere pollutants", "deep learning",

"machine learning", and "artificial intelligence" (see the full search strategy in Supplementary material 1). There were no limitations on language, publication year, or record type. Due to the characteristics of CiteSpace, when collecting data, it is not necessary to exclude literature that appears irrelevant, because this may decrease the sensitivity (or accuracy) of searches, resulting in a lost opportunity to discover new associations (23). The categories and impact factor of the journals can be identified by the Journal Citation Reports of the Web of Science database. Two reviewers independently performed literature retrieval and compared results to ensure accuracy and consistency.

Data analysis

Microsoft Excel 2019 (Redmond WA, USA) was used to conduct data aggregation and analysis. CiteSpace 5.8.R1 software (Drexel University, Philadelphia, PA, USA) was used to analyze the publication characteristics, including record type, publication year, discipline, authors and co-cited authors, countries and institutions, journals, co-cited journals and co-cited references, co-occurrence of keywords and burst keywords. They were identified and described by frequencies. The data was saved in the "Download_ XXX" format and imported into CiteSpace due to the specific data format requirements of that program. The timespan was set as "1990–2021" and slice length was set as 1. The pre-processing of the data mainly removes duplicate documents from the data, and then slices the documents according to time. When words with the same meaning are found in the clustering network, we will add them to the alias list and the alias will be in effect. In addition, the relevant data can also be exported for de-duplication using the Microsoft Excel 2019 (Redmond WA, USA).

Results

A total of 1,835 records were retrieved in this study, including four publication types. Journal articles were the most frequent type (1,386, 75.5%), followed by proceedings (344, 18.7%), reviews (99, 5.4%), and editorials (6, 0.4%). The first publications appeared in 1994, and since then publications have gradually increased, with a surge starting in 2017 (Figure 1).

Distribution of countries/regions and institutions

The records came from 93 countries/regions and 466 institutions. The countries with the highest number of publications were China ($N = 524$), followed by the United States ($N = 455$) (Table 1). In addition, the United States (0.24) and the England (0.27) showed a high degree of centrality (Supplementary material 2). The institution with the highest number of publications was the Chinese Academy of Sciences ($N = 58$), and China accounted for eight of the top 10 research institutions (Table 2). Among the top 10 institutions, only Tsinghua University shows a centrality of >0.1, indicating that it is considered important in its collaborative network.

Distribution of authors and co-cited authors

In total, all records encompassed 610 authors. The top 10 authors and co-cited authors are shown in Table 3. Liu Y was the most prolific author with 15 articles, followed by Guo YM with 13. Six authors had fewer than 10 publications. The highest co-cited authors were Breiman L ($N = 284$), Hochreiter S ($N = 155$) and Liu Y ($N = 152$).

Liu Y is not only the author with the most published articles, but also the author with the top three citations. He is the Chair of the Gangarosa Department of Environmental Health at the Rollins School of Public Health, Emory University, and a visiting professor at the School of Environment, Tsinghua University. His main research interests include the application of satellite remote sensing in air pollution exposure assessment; the potential impact of climate change on population health; Geographic Information System and spatial statistics. Guo YM, the second most published researcher, Li SS, the third most published researcher, and Chen GB, the eighth most published researcher, are part of a research team. They are from the Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine at Monash University, where they conduct research on environmental epidemiology, global environmental change, air pollution and health, exposure assessment, remote sensing modeling, and infectious disease modeling.

Distribution of journals and co-cited journals

The citing journals that published focused on veterinary science, animal science, science, ecology, earth science, and marine biology (see Supplementary material 3). The cited articles focused on systems, computers, environments, toxicology, and nutrition. Of the top 10 co-cited journals, three of these contained articles that have been cited more than 500 times. *Atmospheric Environment* had the highest number of citations ($n = 958$), followed by *Science of the Total Environment* ($n = 742$) (Table 4).
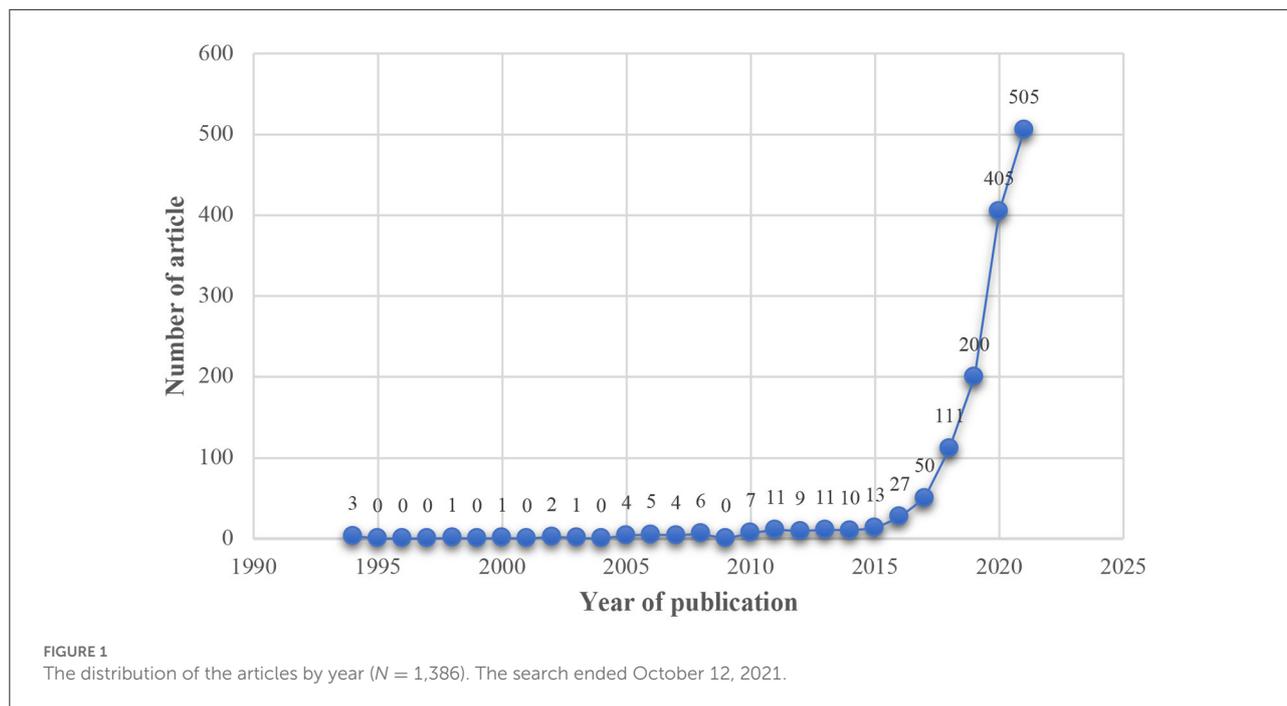
FIGURE 1
The distribution of the articles by year ($N = 1,386$). The search ended October 12, 2021.

TABLE 1 The top 10 countries/regions publishing research on artificial intelligence and air pollution.

Rank	Countries/regions	Count	Centrality
1	China	524	0.08
2	The United States	455	0.24
3	England	136	0.27
4	India	136	0.09
5	South Korea	109	0.04
6	Italy	89	0.14
7	Germany	78	0.05
8	China Taiwan	73	0.06
9	Spain	71	0.12
10	Australia	68	0.19

TABLE 2 The top 10 institutions publishing research on artificial intelligence and air pollution.

Rank	Institutions	Count	Centrality
1	Chinese Academy of Sciences (China)	58	0.06
2	Tsinghua University (China)	33	0.12
3	Wuhan University (China)	31	0.01
4	Peking University (China)	29	0.04
5	Nanjing University of Information Science and Technology (China)	27	0.02
6	Zhejiang University (China)	27	0.04
7	Sun Yat Sen University (China)	25	0.04
8	Emory University (USA)	24	0.02
9	University of Chinese Academy of Sciences (China)	23	0.01
10	National Aeronautics and Space Administration (USA)	22	0.07

Distribution of keywords

The most frequent keywords were machine learning ($N = 462$), air pollution ($N = 410$), and model ($N = 224$) (Supplementary material 4). CiteSpace intelligently classified the research topics into 15 clusters (Figure 2). The modularity (Q) was 0.6023, which was higher than 0.3, indicating that the cluster results were significant. The cluster 0 is the largest cluster (concentration) and cluster 16 is the smallest one (spatial assessment). In recent years, researchers have focused on research about concentration, thermal comfort, air pollution and theoretical prediction (Supplementary material 5). There were eight keywords with the strongest citation bursts, which

identifies hot topics (Figure 3). There were no hot topics identified in 2020 or 2021.

Distribution of references

Among the top 10 co-cited references, one article was co-cited more than 90 times. And only one article was co-cited <50 times (Table 5). The map of reference clustering

(Figure 4) showed that the collaborative network modularity was 0.9388 and the weighted mean silhouette score 0.974, which are both considered very high (17, 18). Similar to the map of keywords clustering (Figure 2, research on concentrations $PM_{2.5}$, was also the focus for reference clustering (Figure 4). Figure 5 shows the top 25 references with the strongest citation bursts. The first reference with citation bursts appeared in 2014, and the last reference with citation bursts appeared in 2019.

TABLE 3 The top 10 authors and co-cited authors publishing research on artificial intelligence and air pollution.

Rank	Authors	Number	Co-cited authors	Co-cited number
1	Liu Y	15	Breiman L	284
2	Guo YM	13	Hochreiter S	155
3	Li SS	11	Liu Y	152
4	Li LF	10	Li X	124
5	Ma J	9	Zhang Y	123
6	Kloog I	9	Di Q	121
7	Fu HB	8	Van Donkelaar A	116
8	Chen GB	8	Hu XF	112
9	Lyapustin A	8	World Health Organization	108
10	Choi Y	7	Lecun Y	107

Discussion

Summary of main results

This study explored research trends on the application of AI in the field of air pollution by using bibliometric analysis and data visualization software. The number of studies on the application of AI in the field of air pollution has risen sharply since 2017. Research institutions in China and the United States were the most active in this field, and most studies were published in journals related to environmental science. The research frontier was concentration prediction, especially about $PM_{2.5}$.

Countries and institutions producing research on artificial intelligence

There may be several reasons for the dominance of Chinese researchers in the field of AI and air pollution. Firstly, on July 8, 2017, China's State Council issued the "New Generation of Artificial Intelligence Development Plan," which sets out the guiding ideology, strategic targets, key tasks and safeguards for the development of China's new generation of AI in 2030. This Plan highlights "intelligent environmental protection"—namely the establishment of an intelligent monitoring big data platform covering the atmosphere, water, and soil. As well, this plan recommends the development of an intelligent

TABLE 4 The top 10 co-cited journals publishing research on artificial intelligence and air pollution.

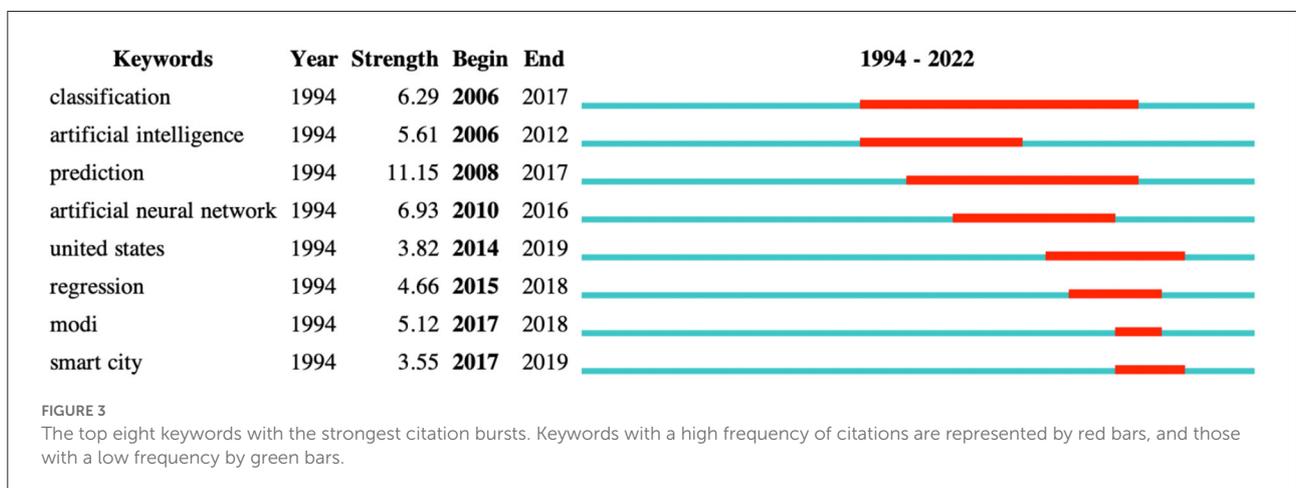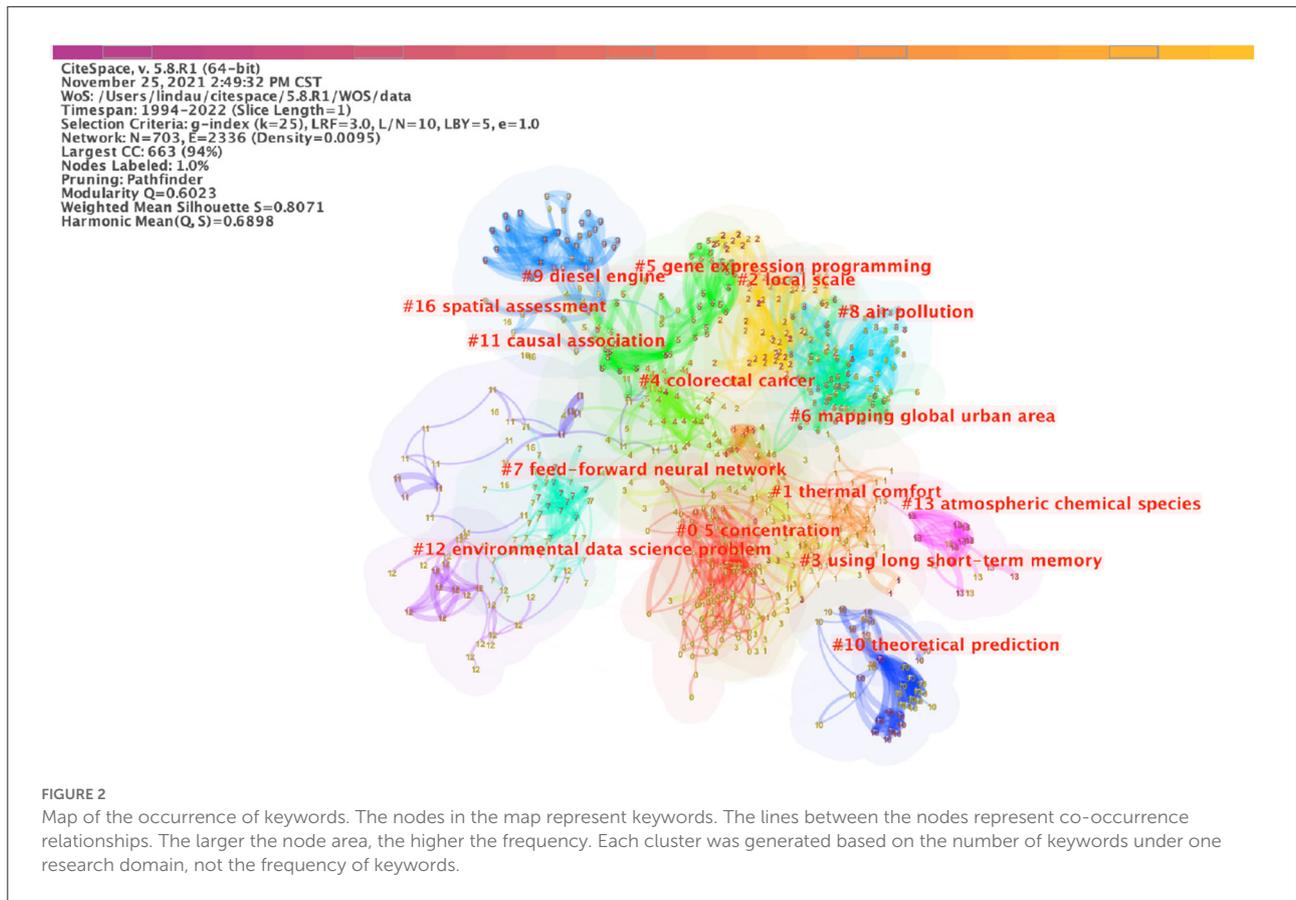| Rank	Co-cited journals	Count	Centrality	Year	IF$_{2020}$	Category
1	Atmospheric Environment	958	0.04	2002	4.798	Meteorology and atmospheric sciences Environmental sciences
2	Science of the Total Environment	742	0.04	2008	7.963	Environmental sciences
3	Environmental Pollution	558	0.03	2008	8.071	Environmental sciences
4	Environmental Science and Technology	460	0.06	1994	9.028	Engineering environmental Environmental sciences
5	Environment International	446	0.05	2008	9.621	Environmental sciences
6	Atmospheric Chemistry and Physics	387	0.03	2014	6.133	Meteorology and atmospheric sciences Environmental sciences
7	Environmental Health Perspectives	370	0.04	2003	9.031	Public, environmental, and occupational health Toxicology Environmental sciences
8	Nature	353	0.03	2001	49.962	Multidisciplinary sciences
9	Environmental Research	345	0.01	2014	6.498	Public, environmental, and occupational health Environmental sciences
10	Machine Learning	339	0.08	2003	2.940	Computer science, artificial intelligence

IF, Impact factor.

FIGURE 2

Map of the occurrence of keywords. The nodes in the map represent keywords. The lines between the nodes represent co-occurrence relationships. The larger the node area, the higher the frequency. Each cluster was generated based on the number of keywords under one research domain, not the frequency of keywords.

FIGURE 3

The top eight keywords with the strongest citation bursts. Keywords with a high frequency of citations are represented by red bars, and those with a low frequency by green bars.

environmental monitoring network and service platform for comprehensive coordination and information sharing. This document also suggests a number of measures, including the development of intelligent prediction models and early warning programs for resource and energy consumption, as well as environmental pollutant emissions. The marked increase in AI-related publications from China since 2017 may reflect the publication of this Plan (24). Secondly, there have been

significant advances in science and technology in recent years in China. The *China Internet Development Report (2021)* released by the Internet Society of China reports that the money invested of the AI industry reached 48.0 billion dollars in 2020, an increase of 15% per year since 2020, a rate slightly higher than the global average growth rate (25). Thirdly, Zhu and Liu predicted that China would overtake the United States as the largest publication producer in 2020 or 2021, if all document

TABLE 5 The Top 10 co-cited references on artificial intelligence and air pollution.

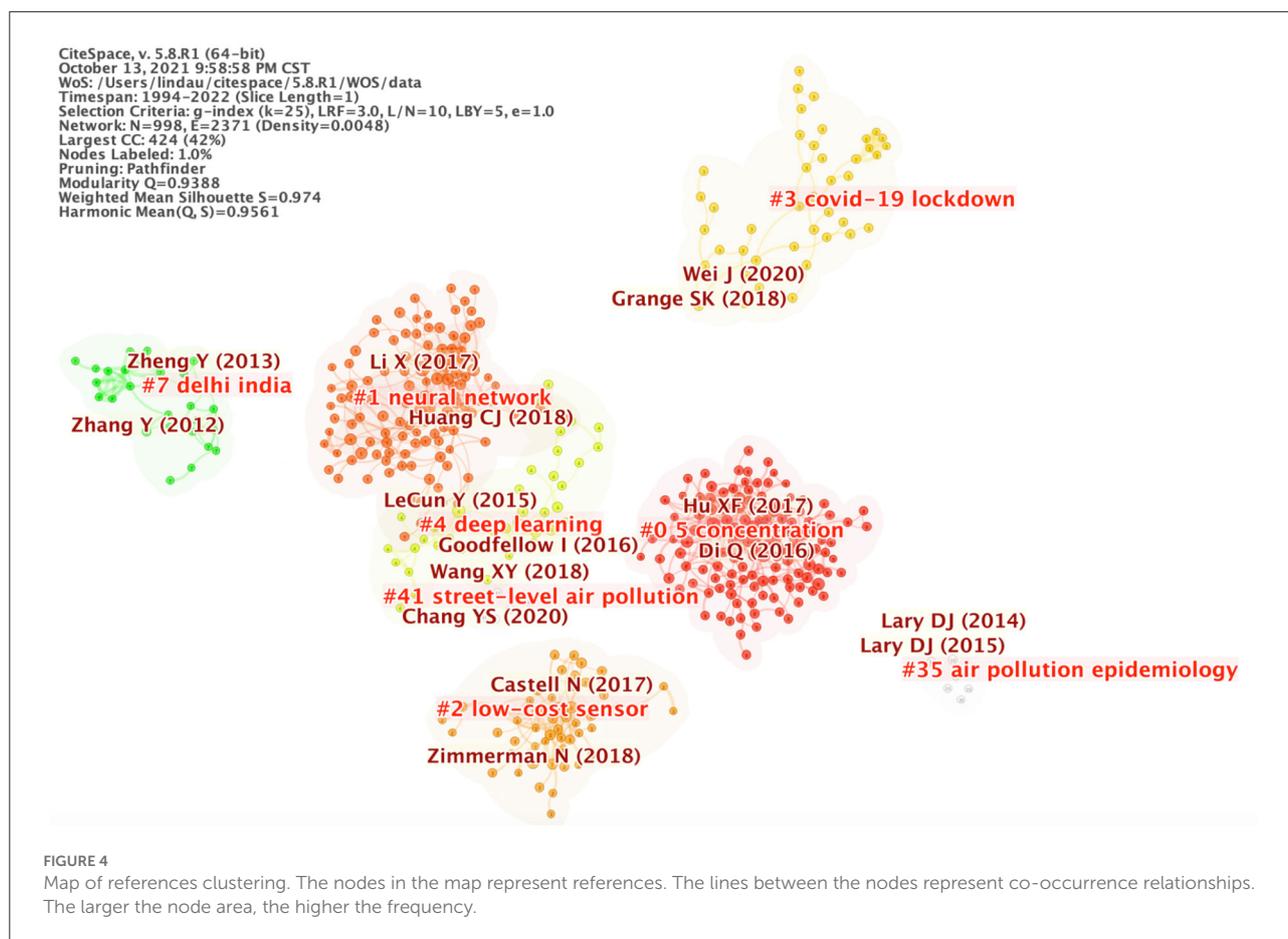| Rank	Title	Journal	Year	First author	First author's affiliation	Centrality	Count
1	Estimating $PM_{2.5}$ concentrations in the conterminous United States using the random forest approach	Environmental Science and Technology	2017	Hu XF	University of Nevada Reno (USA)	0.01	91
2	Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation	Environmental Pollution	2017	Li X	Chinese Academy of Sciences (China)	0.03	77
3	A deep CNN-LSTM Model for particulate matter ($PM_{2.5}$) forecasting in smart cities	Sensors	2018	Huang CJ	Jiangxi University of Science and Technology (China)	0.05	63
4	Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States	Environmental Science and Technology	2016	Di Q	Harvard T.H. Chan School of Public Heath (USA)	0.05	62
5	A machine learning method to estimate $PM_{2.5}$ concentrations across China with remote sensing, meteorological and land use information	Science of the Total Environment	2018	Chen GB	Monash University (Australia)	0.00	61
6	Deep learning	Nature	2015	LeCun Y	Facebook AI Research (USA)	0.02	54
7	Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015	Lancet	2017	Cohen AJ	Health Effects Institute (USA)	0.00	54
8	XGBoost: A scalable tree boosting system	KDD16: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining	2016	Chen TQ	University of Washington (USA)	0.00	53
9	Spatiotemporal prediction of continuous daily $PM_{2.5}$ concentrations across China using a spatially explicit machine learning algorithm	Atmospheric Environment	2017	Zhan Y	Zhejiang University (China)	0.01	51
10	Artificial neural networks forecasting of $PM_{2.5}$ pollution using air mass trajectory based geographic model and wavelet transformation	Atmospheric Environment	2015	Feng X	Peking University (China)	0.02	49

types were considered (26). Given this, it is likely that the increased number of publications in China will have an impact on publications in a variety of fields.

Although China has a high volume of publications, it has a relatively low centrality of 0.08 compared to the United States (0.24) and the England (0.27). Centrality is an index which measures the importance of a node in a network and is mainly used to measure the value of the bridge function of the node in the entire network structure. Generally, nodes >0.1 are considered relatively important (27). This indicates that, despite the high volume of publications from Chinese research institutions, there are few contacts among them, and their connectivity in the overall connectivity is low. In addition, there does not appear to be close cooperation among Chinese research institutions in this field. More extensive cooperation within China and among countries and research institutions would greatly benefit the field of AI in air pollution.

Research authors and journals

This study shows that most of the co-cited journals have an environmental science focus: the highest cited journal is *Atmospheric Environment* which reaches nearly

FIGURE 4
Map of references clustering. The nodes in the map represent references. The lines between the nodes represent co-occurrence relationships. The larger the node area, the higher the frequency.
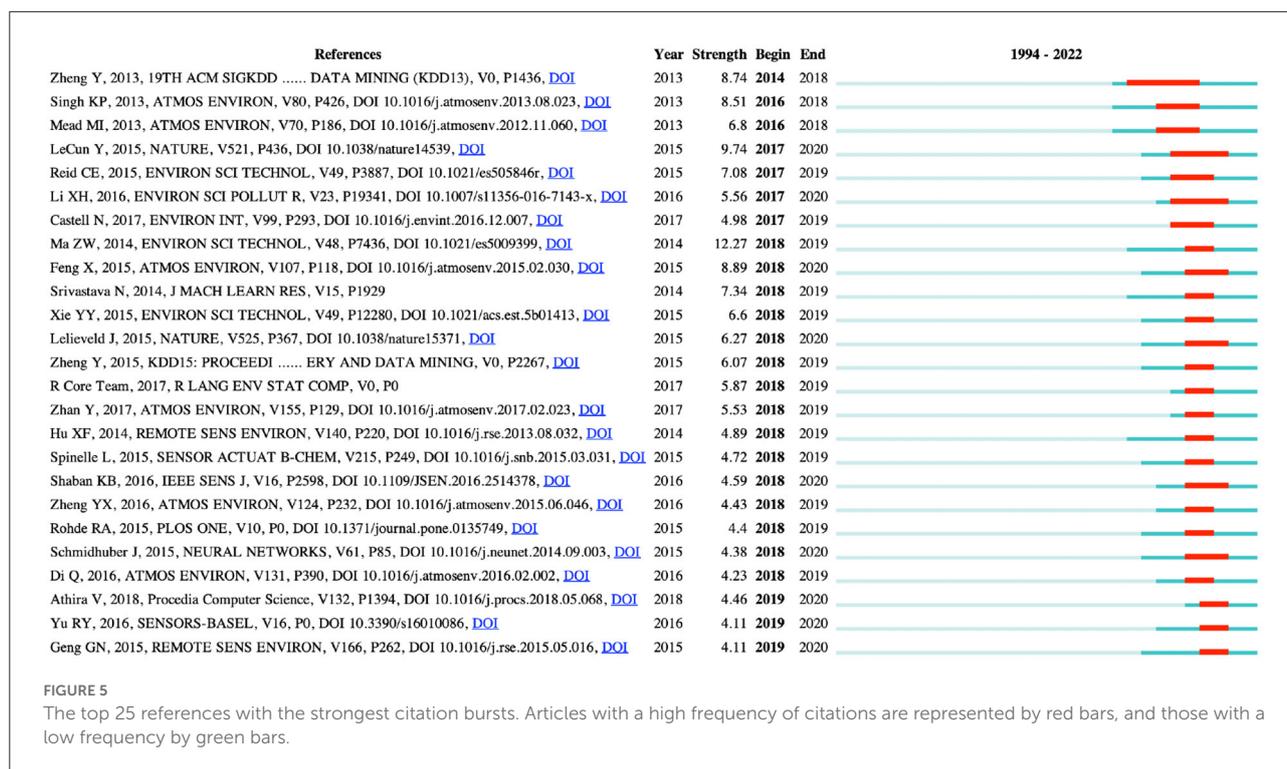
1,000 citations. *Machine Learning* is the only one of the top ten journals which has a singular focus on Computer Science Artificial Intelligence, and the lead researchers mainly belong to the fields of public health and environmental science. This suggests there is a great deal of room for more inter-sectoral research, with input from a broader range of scientists.

Research keywords

The keywords reflect the themes and core content of the research. The most common keywords could be divided into two categories: artificial intelligence (including machine learning, neural network and deep learning) and air pollution (containing air pollution, air quality, exposure and particular matter). The main research theme was prediction of the concentration of air pollution and model analysis through AI technology. Fifteen clusters were obtained with the mapping of keywords. The nodes on the map of timeline viewer of keywords represent one cluster of research and the denser the nodes are the more researchers pay attention to the cluster. Thermal comfort,

which relates to indoor air pollution, is the earliest application and the focus of current research. Thermal comfort refers to the state that most people are satisfied with the objective thermal environment both psychologically and physiologically. Indoor air quality needs to be combined with thermal comfort, where temperature, humidity, and pollutant concentrations are considered in unison and indoor air quality and thermal comfort are inseparable. Initially, traditional mathematical methods were used to give precise definitions of the otherwise vague boundaries between the various levels of thermal sensation. As research progressed, machine learning methods (particularly artificial neural network) were proposed to build algorithmic models for evaluation [28]. Artificial neural networks, which are programmed similarly to the human brain, can process large amounts of data. As data is collected, it is processed by various layers in the program, much like neurons in the human brain, and several types of artificial neural networks are available. AI techniques are now also being used to predict indoor thermal comfort. Moreover, the research on the concentration of air pollutants is also a research hotspot. The future research should be focus on the concentration prediction of air pollution.

References	Year	Strength	Begin	End	1994 - 2022
Zheng Y, 2013, 19TH ACM SIGKDD ...... DATA MINING (KDD13), V0, P1436, DOI	2013	8.74	2014	2018	
Singh KP, 2013, ATMOS ENVIRON, V80, P426, DOI 10.1016/j.atmosenv.2013.08.023, DOI	2013	8.51	2016	2018	
Mead MI, 2013, ATMOS ENVIRON, V70, P186, DOI 10.1016/j.atmosenv.2012.11.060, DOI	2013	6.8	2016	2018	
LeCun Y, 2015, NATURE, V521, P436, DOI 10.1038/nature14539, DOI	2015	9.74	2017	2020	
Reid CE, 2015, ENVIRON SCI TECHNOL, V49, P3887, DOI 10.1021/es505846r, DOI	2015	7.08	2017	2019	
Li XH, 2016, ENVIRON SCI POLLUT R, V23, P19341, DOI 10.1007/s11356-016-7143-x, DOI	2016	5.56	2017	2020	
Castell N, 2017, ENVIRON INT, V99, P293, DOI 10.1016/j.envint.2016.12.007, DOI	2017	4.98	2017	2019	
Ma ZW, 2014, ENVIRON SCI TECHNOL, V48, P7436, DOI 10.1021/es5009399, DOI	2014	12.27	2018	2019	
Feng X, 2015, ATMOS ENVIRON, V107, P118, DOI 10.1016/j.atmosenv.2015.02.030, DOI	2015	8.89	2018	2020	
Srivastava N, 2014, J MACH LEARN RES, V15, P1929	2014	7.34	2018	2019	
Xie YY, 2015, ENVIRON SCI TECHNOL, V49, P12280, DOI 10.1021/acs.est.5b01413, DOI	2015	6.6	2018	2019	
Lelieveld J, 2015, NATURE, V525, P367, DOI 10.1038/nature15371, DOI	2015	6.27	2018	2020	
Zheng Y, 2015, KDD15: PROCEEDI ...... ERY AND DATA MINING, V0, P2267, DOI	2015	6.07	2018	2019	
R Core Team, 2017, R LANG ENV STAT COMP, V0, P0	2017	5.87	2018	2019	
Zhan Y, 2017, ATMOS ENVIRON, V155, P129, DOI 10.1016/j.atmosenv.2017.02.023, DOI	2017	5.53	2018	2019	
Hu XF, 2014, REMOTE SENS ENVIRON, V140, P220, DOI 10.1016/j.rse.2013.08.032, DOI	2014	4.89	2018	2019	
Spinelle L, 2015, SENSOR ACTUAT B-CHEM, V215, P249, DOI 10.1016/j.snb.2015.03.031, DOI	2015	4.72	2018	2019	
Shaban KB, 2016, IEEE SENS J, V16, P2598, DOI 10.1109/JSEN.2016.2514378, DOI	2016	4.59	2018	2020	
Zheng YX, 2016, ATMOS ENVIRON, V124, P232, DOI 10.1016/j.atmosenv.2015.06.046, DOI	2016	4.43	2018	2019	
Rohde RA, 2015, PLOS ONE, V10, P0, DOI 10.1371/journal.pone.0135749, DOI	2015	4.4	2018	2019	
Schmidhuber J, 2015, NEURAL NETWORKS, V61, P85, DOI 10.1016/j.neunet.2014.09.003, DOI	2015	4.38	2018	2020	
Di Q, 2016, ATMOS ENVIRON, V131, P390, DOI 10.1016/j.atmosenv.2016.02.002, DOI	2016	4.23	2018	2019	
Athira V, 2018, Procedia Computer Science, V132, P1394, DOI 10.1016/j.procs.2018.05.068, DOI	2018	4.46	2019	2020	
Yu RY, 2016, SENSORS-BASEL, V16, P0, DOI 10.3390/s16010086, DOI	2016	4.11	2019	2020	
Geng GN, 2015, REMOTE SENS ENVIRON, V166, P262, DOI 10.1016/j.rse.2015.05.016, DOI	2015	4.11	2019	2020	

FIGURE 5
The top 25 references with the strongest citation bursts. Articles with a high frequency of citations are represented by red bars, and those with a low frequency by green bars.

Research references

The co-citation can reflect the researchers' attention (22). According to the citation analysis, seven articles predicted the concentration of $PM_{2.5}$ (16–18, 29–32), one study introduced deep learning (33), one analyzed the global burden resulting from environmental air pollution (34), one described a kind of machine learning, namely Tree boosting (35). These authors of high-cited researches mainly come from the United States and China, similar to the countries publishing the most researches in this field.

Deep learning is based on machine learning algorithms, using multiple layers to progressively extract higher-level features from the raw input (36). Learning can be supervised, semi-supervised or unsupervised (37–40). As the map of references clustering showed, the application of deep learning (particularly artificial neural network) in air pollution has attracted much attention (17, 18). It is also the frontier for researchers to develop models for predicting air pollution [including nitrogen dioxide concentration (41), sulfur dioxide concentration (42), ozone density (43), carbon monoxide concentration (44), and so on (45)]. In the past, research hotspots included the study of air pollution in city streets (46–48), epidemiological studies of the temporal and spatial distribution of air pollution (49–51), and low-cost sensors

for air quality monitoring (52–54). Since 2020, there has been publication of a large volume of literature on the impact of lockdowns due to Coronavirus Disease 2019 (55–60).

Of the 25 references with the strongest citation bursts, eleven articles focused on using AI methods to predict air pollutant concentrations (31–33, 52, 61–67). In five articles, technologies commonly used in environmental science (e.g., satellite-based aerosol optical depth, satellite remote sensing) were combined with AI models to predict air pollutant concentrations (68–72). Four articles investigated the use of low-cost sensors in air quality monitoring (52, 73–75). One article reviewed deep learning in neural networks (38), another examined the status of air pollution in China (76), and yet another computationally analyzed the major outdoor air pollutants and their sources (77). Srivastava and Salakhutdinov showed one specific artificial intelligence technology (78). Li et al. analyzed the distribution of phthalate esters in water and surface sediments of the Pearl River Estuary (79). It further confirms the research hotspot of predicting air pollutant concentrations. Furthermore, low-cost air quality sensors, particularly mobile air quality sensors, can assist people in recognizing the practical needs of air quality monitoring. Unfortunately, there is a lack of advanced technology to provide an accurate measurement of current conditions. Using big data analysis, advanced

algorithms and artificial intelligence techniques, air quality sensors can now provide accurate outdoor air quality analysis and real-time results (80, 81). Lim et al. showed that mobile sampling in conjunction with multiple low-cost air quality monitors could be applied to characterize urban street-level air quality with high spatial resolution, and that machine learning models could further improve model performance (82). Air quality is dynamic and data accuracy can be ensured through calibration, continuous monitoring and air quality data analysis, which is essential to produce valuable and actionable results and recommendations for health decisions in real time.

Limitations

There are several limitations to this study. Firstly, although we used the WOS for our bibliometric analysis, there are other public and commercially available bibliometric databases, such as Scopus. Due to its long history of construction and the fact that it was the only bibliographic database available before Scopus, the WOS database is currently quite well-known on a global scale. Furthermore, the Scopus database covers fewer articles published before 1996, which has some limitations. And only the WOS database was taken into consideration as the data source for this study because the research institution where this research team is based has only purchased access to it. Secondly, the keywords used in the search strategy (e.g., particular matter, sulfur dioxide, and nitrogen dioxide, etc.) related to specific contaminants in the air, which may not lead to the identification of studies in all aspects of air pollution. Thirdly, some top-ranked keywords were uninformative in isolation (e.g., model, pollution, and exposure) and thus could not be analyzed. Therefore, future studies could employ broader search strategies to further explore this literature.

Conclusion

Researches in the field of AI and air pollution are expanding rapidly in recent years. The majority of scholars are from China and the United States. But the United States and the England contribute greatly to the development of the cooperation network. The Chinese Academy of Sciences is the dominant research institution. Cooperation among research institutions appears to be suboptimal, and strengthening cooperation could greatly benefit this field of research. The prediction of air pollutant concentrations, particularly combined use of AI and environmental science method, low-cost air quality sensors, indoor air quality, and thermal comfort are the current research hotspot.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

QG, XS, and YC designed the study. QG collected the data, analyzed the data and wrote the original draft. MR collected the data and reviewed the manuscript. QG, SW, YS, JW, QW, YM, XS, and YC contributed to modifying and reviewing. All authors have read and approved the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.933665/full#supplementary-material

References

1. Ghorani-Azam A, Riahi-Zanjani B, Balali-Mood M. Effects of air pollution on human health and practical measures for prevention in Iran. *J Res Med Sci.* (2016) 21:189646. doi: 10.4103/1735-1995.189646

2. Landrigan PJ. Air pollution and health. *Lancet Public Health.* (2017) 2:e4–5. doi: 10.1016/S2468-2667(16)30023-8

3. Almetwally AA, Bin-Jumah M, Allam AA. Ambient air pollution and its influence on human health and welfare: an overview. *Environ Sci Pollut Res.* (2020) 27:24815–30. doi: 10.1007/s11356-020-09042-2

4. Ye JJ, Wang SS, Fang Y, Zhang XJ, Hu CY. Ambient air pollution exposure and risk of chronic kidney disease: a systematic review of the literature and meta-analysis. *Environ Res.* (2021) 195:110867. doi: 10.1016/j.envres.2021.110867

5. World Health Organization. *COP26 Special Report on Climate Change and Health.* (2021). Available online at: https://www.who.int/publications/i/item/cop26-special-report (accessed January 13, 2022).

6. Wang N, Yan L, Wang Y. Review of theoretical research on artificial intelligence. In: *DEStech Transactions on Computer Science and Engineering, International Conference on Information Technology and Industrial Automation.* Guangzhou (2018). doi: 10.12783/dtcse/iciti2018/29138

7. Chowdhury M, Sadek AW. Advantages and limitations of artificial intelligence. *Artif Intell Applic Crit Transport Issues.* (2012) 6:360–75. Available online at: http://onlinepubs.trb.org/onlinepubs/circulars/ec168.pdf (accessed August 24, 2022).

8. Borana J. Applications of artificial intelligence & associated technologies. In: *Proceeding of International Conference on Emerging Technologies in Engineering, Biomedical, Management and Science.* Jodhpur (2016). Available online at: https://test.globalinfocloud.com/technodigisoftnew/wp-content/uploads/2019/07/Applications-of-Artificial-Intelligence-Associated-Technologies.pdf (accessed August 24, 2022).

9. Noorbakhsh-Sabet N, Zand R, Zhang Y, Abedi V. Artificial intelligence transforms the future of health care. *Am J Med.* (2019) 132:795–801. doi: 10.1016/j.amjmed.2019.01.017

10. Pyayt AL, Mokhov II, Lang B, Krzhizhanovskaya VV, Meijer RJ. Machine learning methods for environmental monitoring and flood protection. *Int J Comput Inf Eng.* (2011) 5:549–54. doi: 10.5281/zenodo.1075060

11. Hino M, Benami E, Brooks N. Machine learning for environmental monitoring. *Nat Sust.* (2018) 1:583–8. doi: 10.1038/s41893-018-0142-9

12. May TO, Livas-García A, Jiménez-Torres M, Cruz May E, López-Manrique LM, Bassam A. Artificial intelligence techniques for modeling indoor building temperature under tropical climate using outdoor environmental monitoring. *J Energy Eng.* (2020) 146:04020004. doi: 10.1061/(ASCE)EY.1943-7897.0000649

13. Masood A, Ahmad K. A review on emerging artificial intelligence (AI) techniques for air pollution forecasting: fundamentals, application and performance. *J Clean Product.* (2021) 322:129072. doi: 10.1016/j.jclepro.2021.129072

14. Hashimoto DA, Witkowski E, Gao L, Meireles O, Rosman G. Artificial intelligence in anesthesiology: current techniques, clinical applications, and limitations. *Anesthesiology.* (2020) 132:379–94. doi: 10.1097/ALN.0000000000002960

15. Bellinger C, Jabbar MSM, Zaïane O, Osornio-Vargas A. A systematic review of data mining and machine learning for air pollution epidemiology. *BMC Public Health.* (2017) 17:907. doi: 10.1186/s12889-017-4914-3

16. Hu X, Belle JH, Meng X, Waller LA, Strickland MJ, Liu Y. Estimating PM2. 5 concentrations in the conterminous United States using the random forest approach. *Environ Sci Technol.* (2017) 51:6936–44. doi: 10.1021/acs.est.7b01210

17. Li X, Peng L, Yao X, Hu Y, You C, Chi T. Long short-term memory neural network for air pollutant concentration predictions: method development and evaluation. *Environ Pollut.* (2017) 231:997–1004. doi: 10.1016/j.envpol.2017.08.114

18. Huang CJ, Kuo PH. A deep cnn-lstm model for particulate matter (PM2. 5) forecasting in smart cities. *Sensors.* (2018) 18:2220. doi: 10.3390/s18072220

19. Chen C. Science mapping: a systematic review of the literature. *J Data Inform Sci.* (2017) 2:1–40. doi: 10.1515/jdis-2017-0006

20. Chen C, Hu Z, Liu S, Tseng H. Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace. *Exp Opin Biol Ther.* (2012) 12:593–608. doi: 10.1517/14712598.2012.674507

21. Braam RR, Moed HF, Van Raan AFJ. Mapping of science by combined co-citation and word analysis. II:

dynamical aspects. *J Am Soc Inform Sci.* (1991) 42:252–66. doi: 10.1002/(SICI)1097-4571(199105)42:4<252::AID-ASI2>3.0.CO;2-G

22. Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J Am Soc Inform Sci Technol.* (2006) 57:359–77. doi: 10.1002/asi.20317

23. Li J, Chen CM. *CiteSpace: Text Mining and Visualization in Scientific Literature.* 2nd Edition. Beijing: Capital University of Economics and Business Press (2016).

24. The State Council the People's Republic of China. *Notice of the State Council on the Issuance of the Development Plan for a New Generation of Artificial Intelligence.* The State Council the People's Republic of China (2017). Available online at: http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm (accessed January 13, 2022).

25. Internet Society of China. *2021 China Internet Conference.* China Internet Development Report 2021. Beijing: Internet Society of China (2021). Available online at: http://www.isc.org.cn/article/40203.html (accessed January 13, 2022).

26. Zhu J, Liu W. Comparing like with like: China ranks first in SCI-indexed research articles since 2018. *Scientometrics.* (2020) 124:1691–700. doi: 10.1007/s11192-020-03525-2

27. Ma D, Yang B, Guan B, Song L, Liu Q, Fan Y, et al. A bibliometric analysis of pyroptosis from 2001 to 2021. *Front Immunol.* (2021) 12:731933. doi: 10.3389/fimmu.2021.731933

28. Gan VJL, Luo H, Tan Y, Deng M, Kwok HL. BIM and data-driven predictive analysis of optimum thermal comfort for indoor environment. *Sensors.* (2021) 21:4401. doi: 10.3390/s21134401

29. Di Q, Kloog I, Koutrakis P, Wang Y, Schwartz J. Assessing PM2. 5 exposures with high spatiotemporal resolution across the continental United States. *Environ Sci Technol.* (2016) 50:4712–21. doi: 10.1021/acs.est.5b06121

30. Chen G, Li S, Knibbs LD, Hamm NA, Cao W, Li T, et al. A machine learning method to estimate PM2. 5 concentrations across China with remote sensing, meteorological and land use information. *Sci Total Environ.* (2018) 636:52–60. doi: 10.1016/j.scitotenv.2018.04.251

31. Zhan Y, Luo Y, Deng X, Chen H, Grieneisen ML, Shen X, et al. Spatiotemporal prediction of continuous daily PM2. 5 concentrations across China using a spatially explicit machine learning algorithm. *Atmosph Environ.* (2017) 155:129–39. doi: 10.1016/j.atmosenv.2017.02.023

32. Feng X, Li Q, Zhu Y, Hou J, Jin L, Wang J. Artificial neural networks forecasting of PM2. 5 pollution using air mass trajectory based geographic model and wavelet transformation. *Atmosph Environ.* (2015) 107:118–28. doi: 10.1016/j.atmosenv.2015.02.030

33. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* (2015) 521:436–44. doi: 10.1038/nature14539

34. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *Lancet.* (2017) 389:1907–18. doi: 10.1016/S0140-6736(17)30505-6

35. Chen T, Guestrin C. Xgboost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM sigkdd International Conference on Knowledge Discovery and Data Mining.* San Francisco, CA (2016). p. 785–94. doi: 10.1145/2939672.2939785

36. Deng L, Yu D. Deep learning: methods and applications. *Found Trends Signal Process.* (2014) 7:197–387. doi: 10.1561/2000000039

37. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Trans Patt Anal Mach Intell.* (2013) 35:1798–828. doi: 10.1109/TPAMI.2013.50

38. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw.* (2015) 61:85–117. doi: 10.1016/j.neunet.2014.09.003

39. Lauzon FQ. An introduction to deep learning. In: *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA).* Montreal, QC: IEEE (2012). p. 1438–9. doi: 10.1109/ISSPA.2012.6310529

40. Goodfellow IJ, Courville A, Bengio Y. Scaling up spike-and-slab models for unsupervised feature learning. *IEEE Trans Patt Anal Mach Intell.* (2012) 35:1902–14. doi: 10.1109/TPAMI.2012.273

41. Shams SR, Jahani A, Kalantary S, Moeinaddini M, Khorasani N. Artificial intelligence accuracy assessment in NO2 concentration forecasting of metropolises air. *Sci Rep.* (2021) 11:1805. doi: 10.1038/s41598-021-81455-6

42. Shams SR, Jahani A, Kalantary S, Moeinaddini M, Khorasani N. The evaluation on artificial neural networks (ANN) and multiple linear

regressions (MLR) models for predicting SO2 concentration. *Urban Clim.* (2021) 37:100837. doi: 10.1016/j.uclim.2021.100837

43. Shams SR, Jahani A, Moeinaddini M, Khorasani N, Kalantary S. Forecasting ozone density in tehran air using a smart data-driven approach. *J Health Saf Work.* (2020) 10:406–20. Available online at: http://jhsw.tums.ac.ir/article-1-6415-en.html (accessed August 24, 2022).

44. Shams SR, Jahani A, Moeinaddini M, Khorasani N. Air carbon monoxide forecasting using an artificial neural network in comparison with multiple regression. *Model Earth Syst Environ.* (2020) 6:1467–5. doi: 10.1007/s40808-020-00762-5

45. Chang YS, Abimannan S, Chiao HT, Lin CY, Huang YP. An ensemble learning based hybrid model and framework for air pollution forecasting. *Environ Sci Pollut Res.* (2020) 27:38155–68. doi: 10.1007/s11356-020-09855-1

46. Boogaard H, Janssen NAH, Fischer PH, Kos GP, Weijers EP, Cassee FR, et al. Impact of low emission zones and local traffic policies on ambient air pollution concentrations. *Sci Total Environ.* (2012) 435:132–40. doi: 10.1016/j.scitotenv.2012.06.089

47. Huang Y, Lei C, Liu CH, Perez P, Forehead H, Kong S, et al. A review of strategies for mitigating roadside air pollution in urban street canyons. *Environ Pollut.* (2021) 280:116971. doi: 10.1016/j.envpol.2021.116971

48. Vardoulakis S, Kettle R, Cosford P, Lincoln P, Holgate S, Grigg J, et al. Local action on outdoor air pollution to improve public health. *Int J Public Health.* (2018) 63:557–65. doi: 10.1007/s00038-018-1104-8

49. Lary DJ, Faruque FS, Malakar N, Moore A, Roscoe B, Adams ZL, et al. Estimating the global abundance of ground level presence of particulate matter (PM2.5). *Geospat Health.* (2014) 8:S611–30. doi: 10.4081/gh.2014.292

50. Wu D, Zewdie GK, Liu X, Kneen MA, Lary DJ. Insights into the morphology of the East Asia PM2. 5 annual cycle provided by machine learning. *Environ Health Insights.* (2017) 11:1178630217699611. doi: 10.1177/1178630217699611

51. Wu D, Lary DJ, Zewdie GK, Liu X. Using machine learning to understand the temporal morphology of the PM 2.5 annual cycle in East Asia. *Environ Monit Assess.* (2019) 191:1–14. doi: 10.1007/s10661-019-7424-1

52. Castell N, Dauge FR, Schneider P, Vogt M, Lerner U, Fishbain B, et al. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environ Int.* (2017) 99:293–302. doi: 10.1016/j.envint.2016.12.007

53. Schneider P, Castell N, Vogt M, Dauge FR, Lahoz WA, Bartonova A. Mapping urban air quality in near real-time using observations from low-cost sensors and model information. *Environ Int.* (2017) 106:234–47. doi: 10.1016/j.envint.2017.05.005

54. Castell N, Schneider P, Grossberndt S, Fredriksen MF, Sousa-Santos G, Vogt M, et al. Localized real-time information on outdoor air quality at kindergartens in Oslo, Norway using low-cost sensor nodes. *Environ Res.* (2018) 165:410–9. doi: 10.1016/j.envres.2017.10.019

55. World Health Organization. *WHO Director-General's Remarks at the Media Briefing on 2019-nCoV on 11 February 2020.* World Health Organization (2020). Available online at: https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-nCoV-on-11-february-2020 (accessed January 13, 2022).

56. Yang HY, Duan GC. Analysis on the epidemic factors for COVID-19. *Chin J Prevent Medline.* (2020) 54:608–13. doi: 10.3760/cma.j.cn112150-20200227-00196

57. World Health Organization. *WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 11 March 2020.* World Health Organization (2020). Available online at: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-\$-\$11-march-2020 (accessed January 13, 2022).

58. Jephcote C, Hansell AL, Adams K, Gulliver J. Changes in air quality during COVID-19 'lockdown' in the United Kingdom. *Environ Pollut.* (2021) 272:116011. doi: 10.1016/j.envpol.2020.116011

59. Hua J, Zhang Y, de Foy B, Gulliver J. Quantitative estimation of meteorological impacts and the COVID-19 lockdown reductions on NO2 and PM2.5 over the Beijing area using generalized additive models (GAM). *J Environ Manage.* (2021) 291:112676. doi: 10.1016/j.jenvman.2021.112676

60. Fu S, Guo M, Fan L, Deng Q, Han D, Wei Y, et al. Ozone pollution mitigation in guangxi (south China) driven by meteorology and anthropogenic emissions during the COVID-19 lockdown. *Environ Pollut.* (2021) 272:115927. doi: 10.1016/j.envpol.2020.115927

61. Singh KP, Gupta S, Rai P. Identifying pollution sources and predicting urban air quality using ensemble learning methods. *Atmosph Environ.* (2013) 80:426–37. doi: 10.1016/j.atmosenv.2013.08.023

62. Reid CE, Jerrett M, Petersen ML, Pfister GG, Morefield PE, Tager IB, et al. Spatiotemporal prediction of fine particulate matter during the 2008 Northern

California wildfires using machine learning. *Environ Sci Technol.* (2015) 49:3887–96. doi: 10.1021/es505846r

63. Di Q, Koutrakis P, Schwartz J. A hybrid prediction model for PM2. 5 mass and components using a chemical transport model and land use regression. *Atmosph Environ.* (2016) 131:390–9. doi: 10.1016/j.atmosenv.2016.02.002

64. Yu R, Yang Y, Yang L, Han G. RAQ–A random forest approach for predicting air quality in urban sensing systems. *Sensors.* (2016) 16:86. doi: 10.3390/s16010086

65. Athira V, Geetha P, Vinayakumar R, Soman KP. Deepairnet: applying recurrent networks for air quality prediction. *Proc Comp Sci.* (2018) 132:1394–403. doi: 10.1016/j.procs.2018.05.068

66. Zheng Y, Liu F, Hsieh HP. U-air: when urban air quality inference meets big data. In: *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* Chicago, IL (2013). p. 1436–44. doi: 10.1145/2487575.2488188

67. Zheng Y, Yi X, Li M, Li R, Shan Z, Chang E, et al. Forecasting fine-grained air quality based on big data. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* Sydney, NSW (2015). p. 2267–76. doi: 10.1145/2783258.2788573

68. Ma ZW, Hu X, Huang L, Bi J, Liu Y. Estimating ground-level PM2. 5 in China using satellite remote sensing. *Environ Sci Technol.* (2014) 48:7436–44. doi: 10.1021/es5009399

69. Geng G, Zhang Q, Martin RV, van Donkelaar A, Huo H, Che H, et al. Estimating long-term PM2. 5 concentrations in China using satellite-based aerosol optical depth and a chemical transport model. *Remote Sens Environ.* (2015) 166:262–70. doi: 10.1016/j.rse.2015.05.016

70. Xie Y, Wang Y, Zhang K, Dong W, Lv B, Bai Y. Daily estimation of ground-level PM2. 5 concentrations over Beijing using 3 km resolution MODIS AOD. *Environ Sci Technol.* (2015) 49:12280–8. doi: 10.1021/acs.est.5b01413

71. Hu X, Waller LA, Lyapustin A, Wang Y, Al-Hamdan MZ, Crosson WL, et al. Estimating ground-level PM2. 5 concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote Sens Environ.* (2014) 140:220–32. doi: 10.1016/j.rse.2013.08.032

72. Zheng Y, Zhang Q, Liu Y, Geng G, He K. Estimating ground-level PM2. 5 concentrations over three megalopolises in China using satellite-derived aerosol optical depth measurements. *Atmosph Environ.* (2016) 124:232–42. doi: 10.1016/j.atmosenv.2015.06.046

73. Mead MI, Popoola OAM, Stewart GB, Landshoff P, Calleja M, Hayes M, et al. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmosph Environ.* (2013) 70:186–203. doi: 10.1016/j.atmosenv.2012.11.060

74. Spinelle L, Gerboles M, Villani MG, Aleixandre M, Bonavitacola F. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: ozone and nitrogen dioxide. *Sens Actuators B Chem.* (2015) 215:249–57. doi: 10.1016/j.snb.2015.03.031

75. Shaban KB, Kadri A, Rezk E. Urban air pollution monitoring system with forecasting models. *IEEE Sens J.* (2016) 16:2598–606. doi: 10.1109/JSEN.2016.2514378

76. Rohde RA, Muller RA. Air pollution in China: mapping of concentrations and sources. *PLoS ONE.* (2015) 10:e0135749. doi: 10.1371/journal.pone.0135749

77. Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature.* (2015) 525:367–71. doi: 10.1038/nature15371

78. Srivastava N, Salakhutdinov R. Multimodal learning with deep boltzmann machines. *J Mach Learn Res.* (2014) 15:2949–80. Available online at: https://jmlr.org/papers/volume15/srivastava14b/srivastava14b.pdf (accessed August 24, 2022).

79. Li X, Yin P, Zhao L. Phthalate esters in water and surface sediments of the Pearl River Estuary: distribution, ecological, and human health risks. *Environ Sci Pollut Res.* (2016) 23:19341–9. doi: 10.1007/s11356-016-7143-x

80. Vajs I, Drajic D, Gligoric N, Radovanovic I, Popovic I. Developing relative humidity and temperature corrections for low-cost sensors using machine learning. *Sensors.* (2021) 21:3338. doi: 10.3390/s21103338

81. Singh D, Dahiya M, Kumar R, Nanda C. Sensors and systems for air quality assessment monitoring and management: a review. *J Environ Manage.* (2021) 289:112510. doi: 10.1016/j.jenvman.2021.112510

82. Lim CC, Kim H, Vilcassim MJR, Thurston GD, Gordon T, Chen LC, et al. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea. *Environ Int.* (2019) 131:105022. doi: 10.1016/j.envint.2019.105022