sustainability

MDPI

Review

Towards Federated Learning and Multi-Access Edge Computing for Air Quality Monitoring: Literature Review and Assessment

Satheesh Abimannan [1,*,†], El-Sayed M. El-Alfy [2,†], Shahid Hussain [3,*], Yue-Shan Chang [4], Saurabh Shukla [5], Dhivyadharsini Satheesh [6] and John G. Breslin [7]

1 Amity School of Engineering and Technology, Amity University Maharashtra, Mumbai 410206, India
2 Fellow SDAIA-KFUPM Joint Research Center for Artificial Intelligence, Interdisciplinary Research Center of Intelligent Secure Systems, Information and Computer Science Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia; alfy@kfupm.edu.sa
3 Innovation Value Institute (IVI), School of Business, National University of Ireland Maynooth (NUIM), W23 F2H6 Maynooth, Ireland
4 Department of Computer Science and Information Engineering, National Taipei University, Taipei 10608, Taiwan; ysc@mail.ntpu.edu.tw
5 Department of Computer Science (CS), Indian Institute of Information Technology, Lucknow (IIIT L), Lucknow 226002, India; saurabh.shukla@iiitl.ac.in
6 School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology (VIT), Vellore 632014, India; dhivyadharsini.sk2021@vitstudent.ac.in
7 Department of Electrical and Electronic Engineering, Data Science Institute, University of Galway, H91 TK33 Galway, Ireland; john.breslin@nuigalway.ie
* Correspondence: sabimannan@mum.amity.edu (S.A.); shahid.hussain@mu.ie (S.H.)
† These authors contributed equally to this work.

Abstract: Systems for monitoring air quality are essential for reducing the negative consequences of air pollution, but creating real-time systems encounters several challenges. The accuracy and effectiveness of these systems can be greatly improved by integrating federated learning and multi-access edge computing (MEC) technology. This paper critically reviews the state-of-the-art methodologies for federated learning and MEC-enabled air quality monitoring systems. It discusses the immense benefits of federated learning, including privacy-preserving model training, and MEC, such as reduced latency and improved response times, for air quality monitoring applications. Additionally, it highlights the challenges and requirements for developing and implementing real-time air quality monitoring systems, such as data quality, security, and privacy, as well as the need for interpretable and explainable AI-powered models. By leveraging such advanced techniques and technologies, air monitoring systems can overcome various challenges and deliver accurate, reliable, and timely air quality predictions. Moreover, this article provides an in-depth analysis and assessment of the state-of-the-art techniques and emphasizes the need for further research to develop more practical and affordable AI-powered decentralized systems with improved performance and data quality and security while ensuring the ethical and responsible use of the data to support informed decision making and promote sustainability.

Keywords: federated learning; multi-access edge computing; air quality monitoring; climate change; privacy-preserving methods; sustainable urban environments

check for updates

Citation: Abimannan, S.; El-Alfy, E.-S.M.; Hussain, S.; Chang, Y.-S.; Shukla, S.; Satheesh, D.; Breslin, J.G. Towards Federated Learning and Multi-Access Edge Computing for Air Quality Monitoring: Literature Review and Assessment. *Sustainability* **2023**, *15*, 13951. https://doi.org/10.3390/su151813951

Academic Editors: Alan Randall and Ying Sheng

Received: 4 July 2023
Revised: 5 September 2023
Accepted: 14 September 2023
Published: 20 September 2023

1. Introduction

With the current technological advances and the concerns raised about pollution and climate change in modern societies, it is essential to monitor both the interior and outdoor air quality because the rising levels of air pollutants and toxins present a serious risk to both human health and the environment. For instance, the ambient air pollution is estimated to cause 4.2 million premature deaths globally each year, making it one of the leading causes of death worldwide (https://www.who.int/news-room/fact-sheets/detail/

ambient-(outdoor)-air-quality-and-health (accessed on 5 January 2023)). Air pollutants cause contamination or changes in the natural characteristics of indoor and outdoor environments through various hazardous physical, chemical, and biological substances; e.g., particulate matter (PM_{10}, $PM_{2.5}$), carbon monoxide (CO), nitrogen oxides (NO_x), sulfur dioxide (SO_2), ammonia (NH_3), tropospheric (ground-level) ozone (O_3), and air toxics. They can cause adverse health problems, such as respiratory and cardiovascular diseases, cancer, and early mortality [1–3]. Air quality monitoring systems play a crucial role in identifying the sources of pollutants in various contexts (e.g., indoor, outdoor, rural, urban, industrial, etc.), measuring their concentrations in the atmosphere, and assessing the potential health risks that the public is exposed to. By implementing targeted mitigation measures, these systems also assist in reducing the negative effects of air pollution on the ecosystem.

Recently, there has been significant success in addressing air pollution and related climate change issues through a variety of methods, from policies to technical shifts in industry, transportation, waste management, and renewable energy production [4–6]. However, the development of accurate, reliable, scalable, and cost-effective air quality monitoring systems is a challenging task [7–11]. The design criteria should take into consideration the complex nature of measuring air pollution, the need for real-time monitoring, and the high cost of deploying and maintaining the monitoring equipment. Moreover, air quality monitoring is also important to identify areas with high levels of pollution and assess the effectiveness of mitigation strategies, such as emission control measures, urban planning, and transportation policies, implemented to reduce pollution levels [6,12]. Current air quality monitoring systems use stationary sensors, which have limited coverage and can be impacted by elements such as wind directions and weather conditions [13–19]. Although these stationary sensors are essential for measuring pollutant concentrations, they represent only one aspect of air quality monitoring. To gain a comprehensive understanding of air quality dynamics, it is essential to incorporate data analysis into the monitoring process. Additionally, data from stationary sensors, along with other sources, such as weather stations, satellite observations, social media, and mobile devices, provide valuable information for air quality analysis, forecasting, and decision making. Therefore, a holistic smart platform that considers both the measurement equipment and the data analysis process is necessary to advance air quality monitoring systems.

AI-powered systems in air quality monitoring have a variety of possible uses. For instance, they can be used for air quality forecasting [20,21], source identification [22], anomaly detection [23,24], fault diagnosis [25], event detection [26], air pollution control in ITS [27], exposure assessment [28], environmental and health impact assessment [29–31], and air quality monitoring network optimization [13]. These applications have the potential to improve the accuracy and efficiency of air quality monitoring and forecasting systems. Lately, several suggestions have evolved based on mobile edge computing or multi-access edge computing (MEC), which was previously known as mobile edge computing but then was generalized and standardized by the European Telecommunications Standards Institute (ETSI) (https://www.etsi.org/technologies/multi-access-edge-computing (accessed on 5 January 2023)), to promote digital transformation and low-latency connectivity for urban informatics and environmental monitoring [7,32–35]. MEC integrates mobile cloud–edge computing to enable distributed computation offloading and real-time data processing and analysis close to the data source [36,37]. Empowered with federated learning (FL), which is a recently proposed decentralized machine learning framework, multiple heterogeneous edge nodes can collaboratively train a shared global model without sharing the raw data and while preserving data privacy and reducing the bandwidth required for data transfer [38–47]. FL algorithms, such as federated averaging and federated stochastic gradient descent, have been successfully applied in air quality monitoring [47]. Various strategies have been proposed in the literature to combine MEC and FL, reducing reliance on centralized systems. Solutions can be designed based on "AirSense" (https://www.energomonitor.com/airsense/ (accessed on 5 January

2023)), "SmartAir" (https://smartairfilters.com/en/blog/category/air-quality-monitors/ (accessed on 5 January 2023)), or the Highly Accurate and Autonomous Programmable Platform for Providing Air Pollution Data Services to Drivers and the Public (HAPADS) (https://www.nilu.com/2021/07/hapads-a-mobile-air-quality-monitoring-platform/ (accessed on 5 January 2023)) to use edge computing and federated learning for real-time air quality measurements.

The main goal of this paper is to provide a review and thorough analysis of the state-of-the-art federated learning frameworks and how they operate with multi-access edge computing for intelligent air quality monitoring and forecasting. To accomplish this, we adopted a methodology that included the use of recognized academic databases, inclusion criteria, screening procedures, and quality assessment to ensure the literature review was objective and concentrated on the most pertinent studies. We then synthesized the information extracted from each study to provide a summary of the current state of research in the field, identify recurring themes and trends, and highlight gaps where additional research is needed. Thus, the findings from the analysis can function as a trustworthy source with insightful information for researchers and practitioners in the field. Additionally, the paper explores prospective research avenues and potential solutions for addressing various design requirements and challenges in developing and implementing real-time air quality monitoring systems.

The paper is organized as follows. Section 2 describes the methodology followed in this paper. Section 3 provides an overview of edge computing and edge intelligence and their potential applications in monitoring and forecasting air quality. Section 4 presents a discussion of the design requirements and challenges for developing and implementing real-time air quality monitoring systems. Section 5 discusses future research directions and potential solutions. Finally, the paper concludes in Section 6.

2. Search and Screening Methodology

We began with a systematic search to collect related literature on federated learning and multi-access edge computing for air quality monitoring. Finding pertinent academic databases that would provide a wide range of research articles on the topic was the first step in this literature study.

The search for relevant data and information was conducted using the following digital databases: Mendeley (https://www.mendeley.com/ (accessed on 9 January 2023)), IEEE Xplore Digital Library (http://ieeexplore.ieee.org (accessed on 9 January 2023)), ACM Digital Library (http://dl.acm.org (accessed on 9 January 2023)), Science Direct (http://www.sciencedirect.com (accessed on 9 January 2023)), Scopus (https://www.elsevier.com/solutions/scopus (accessed on 9 January 2023)), and Google Scholar (https://scholar.google.com/) (accessed on 9 January 2023)). These libraries are widely recognized as reputable and reliable sources of academic literature.

We defined various related keywords and search criteria, as shown in Table 1, and then we carried out an intensive search in the chosen databases. We used various combinations of related terms and ended up with more than 60,000 documents, including duplicates. After limiting the search to the title and abstract and removing duplicates, we found around 2000 publications. We conducted a detailed screening to remove irrelevant documents and augmented the remaining ones with some papers manually found to provide an overview of related technologies and machine learning approaches. The final list had 175 publications, which were distributed over various years, as shown in Figure 1, and of various types, as shown in Figure 2. We then analyzed the distribution of included references under each of the categories in Table 1 and the results are shown in Table 2 and Figure 3.

Table 1. Categories and search criteria.

Category	Related Terms	Count
C1: Air quality-related	Air quality index (AQI), air pollution, aerosols or particulate matter (PMx), NOx (nitric oxide, nitrogen dioxide, nitrogen oxides), COx (carbon monoxide, carbon dioxide), ammonia (NH_3), ozone (O_3), toxics, air contamination, emissions of vehicles and plants, volatile organic compounds, radioactive pollution, thermal pollution, outdoor/ambient air quality, indoor air quality	117
C2: Domain	Health effects, environment impacts, smart city, public safety, security, simulation	37
C3: AI-related	Artificial intelligence (AI), machine learning (ML), deep learning (DL), reinforcement learning (RL), federated learning (FL), data mining, big data, regression, forecasting, fuzzy control, multivariate time series, anomaly detection	133
C4: System architecture	Internet of Things (IoT, IIoT, IoE, IoX), cloud computing, cloud–edge computing, fog computing, edge computing, sensor network, multi-access edge computing (MEC), mobile edge computing, computation offloading, federated learning, edge intelligence, 5G edge computing	136
C5: Federated learning-related	Federated learning, computation offloading, edge intelligence	78
C6: MEC-related	Multi-access edge computing, MEC, mobile edge computing	65
C7	C1 AND C5 AND C6	39
C8	C1 AND C3 AND C4	72

Table 2. Distribution of included publications in terms of years of publication for each included category.

Category	1993	2000	2002	2010	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	Total
C1	1	1	1	2	1	2	1	6	9	19	28	34	17	2	124
C2	1	1	1	2	1	4	2	3	3	15	14	20	8	1	76
C3		1		1	1	2	3	5	7	21	31	41	22	1	136
C4						3	5	7	11	22	32	38	20		138
C5						2	1	4	5	14	17	24	11		78
C6						1		4	4	14	16	19	8		66
C7								4	3	6	8	13	5		39
C8						1	1	4	6	8	20	24	9		73

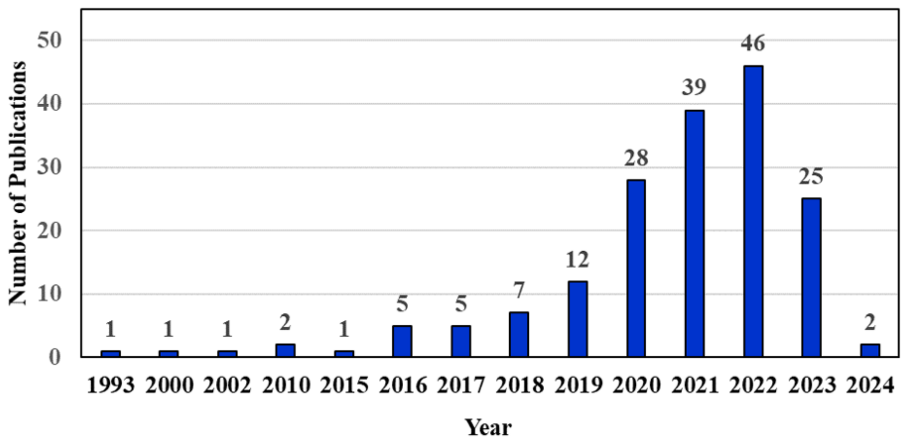


Figure 1. Distribution of selected references per year.

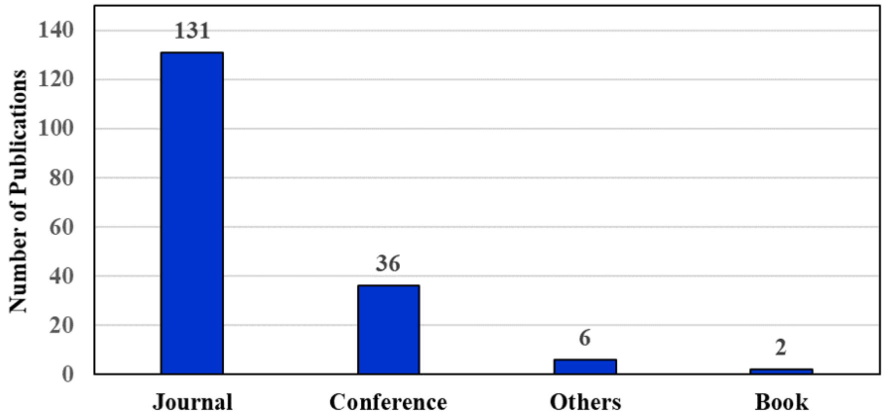


Figure 2. Distribution of selected references based on the type of publication.

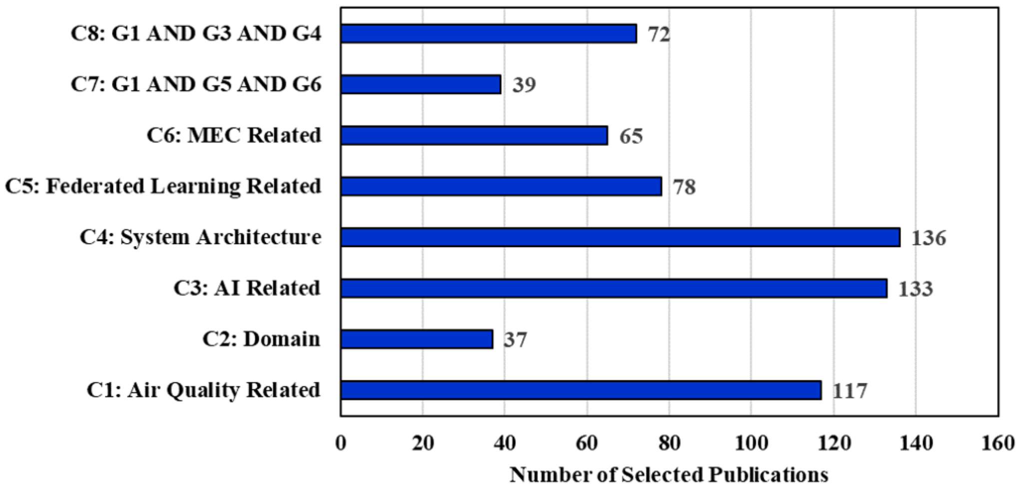


Figure 3. Distribution of selected references per category.

The search results were meticulously screened to eliminate any duplicate or irrelevant studies. The screening process encompassed a careful review of the titles, abstracts, and keywords of the articles to guarantee that they fulfilled the requirements for inclusion in this review paper. Studies that failed to meet the inclusion criteria were excluded from the analysis. This critical screening process aided in focusing the literature review on the most pertinent studies while ensuring that all studies included in the analysis were of high caliber.

Following the initial screening process, the chosen studies were subjected to detailed reading and analysis to extract relevant information, such as the research method, findings, and limitations of each study. The extracted information was synthesized to provide an overview of the current state of research in the field, identify common themes and trends in research, and highlight areas where further research is needed. This synthesis process helped to ensure that the literature review was comprehensive, informative, and provided an in-depth analysis of the state of the art of federated learning and multi-access edge computing in air quality monitoring and forecasting.

3. Edge Computing and Intelligence for Air Quality Monitoring and Predictions

The rise of edge computing and edge intelligence has stimulated advancements in air quality monitoring and forecasting [4,48]. Edge computing offers several operational advantages [47,49]. Primarily, it brings data processing and analysis closer to the source of data generation, bypassing the need for centralized cloud servers. This proximity aids in producing real-time insights into air quality, enhancing the efficiency of environmental monitoring [50–52]. As illustrated in Figure 4, data collection encompasses a variety of sources, such as stationary and mobile air quality IoT sensors, meteorological stations, satellite remote sensing (https://airquality.gsfc.nasa.gov/ (accessed on 25 January 2023)), and social media platforms [53–58]. These sources offer a plethora of information related to air quality, the spatial distribution of greenhouse gases, and the factors influencing air quality, such as traffic flow, meteorological conditions, and human activities [47,59]. However, edge computing also brings new challenges for data collection, integration, and management due to the heterogeneity of sensing devices used to measure the various types of pollutants and the related social, meteorological, and ambient conditions. Depending on the ultimate objective and budget, data can be sampled at different locations and in different time periods with a sensing system deployed with high density [60].

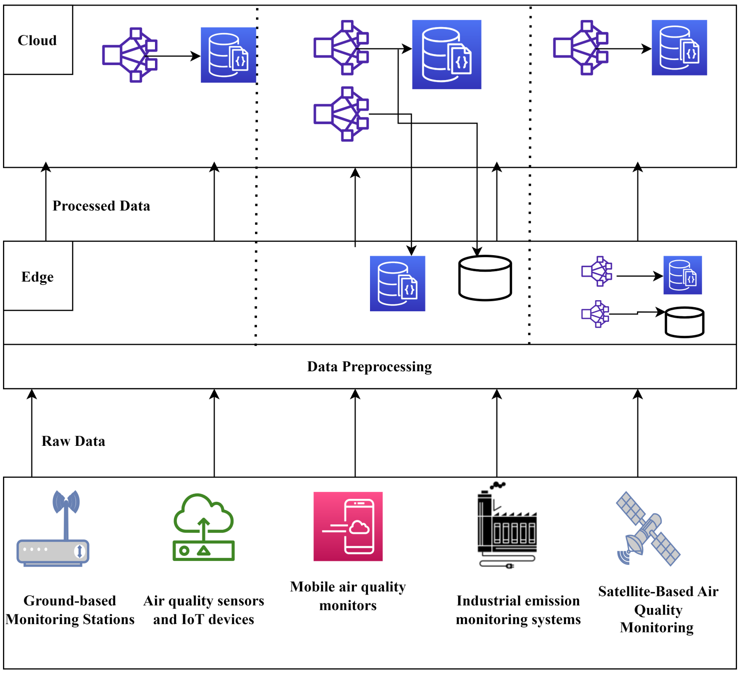


Figure 4. Air quality monitoring system supported by Internet of Things for data collection and processing [61].

Figure 5 depicts the main steps in harnessing various data sources and applying advanced artificial intelligence techniques to monitor and predict air quality. It highlights the crucial role that a combination of diverse data sources and advanced AI techniques can play in creating reliable and useful models for predicting and managing air quality. Initially, the system embarks on an exhaustive data collection phase, wherein it harnesses a variety of environmental parameters. This is accomplished using numerous types of sensors, which monitor aspects such as air quality, temperature, and humidity. Additionally, the system takes advantage of satellite technology with remote sensing data, complementing this with meteorological data that include parameters like wind speed and precipitation, as well as outputs from other meteorological models. Such a broad spectrum of data sources bolsters the comprehensiveness and reliability of the input, thus laying a strong foundation for the subsequent steps.

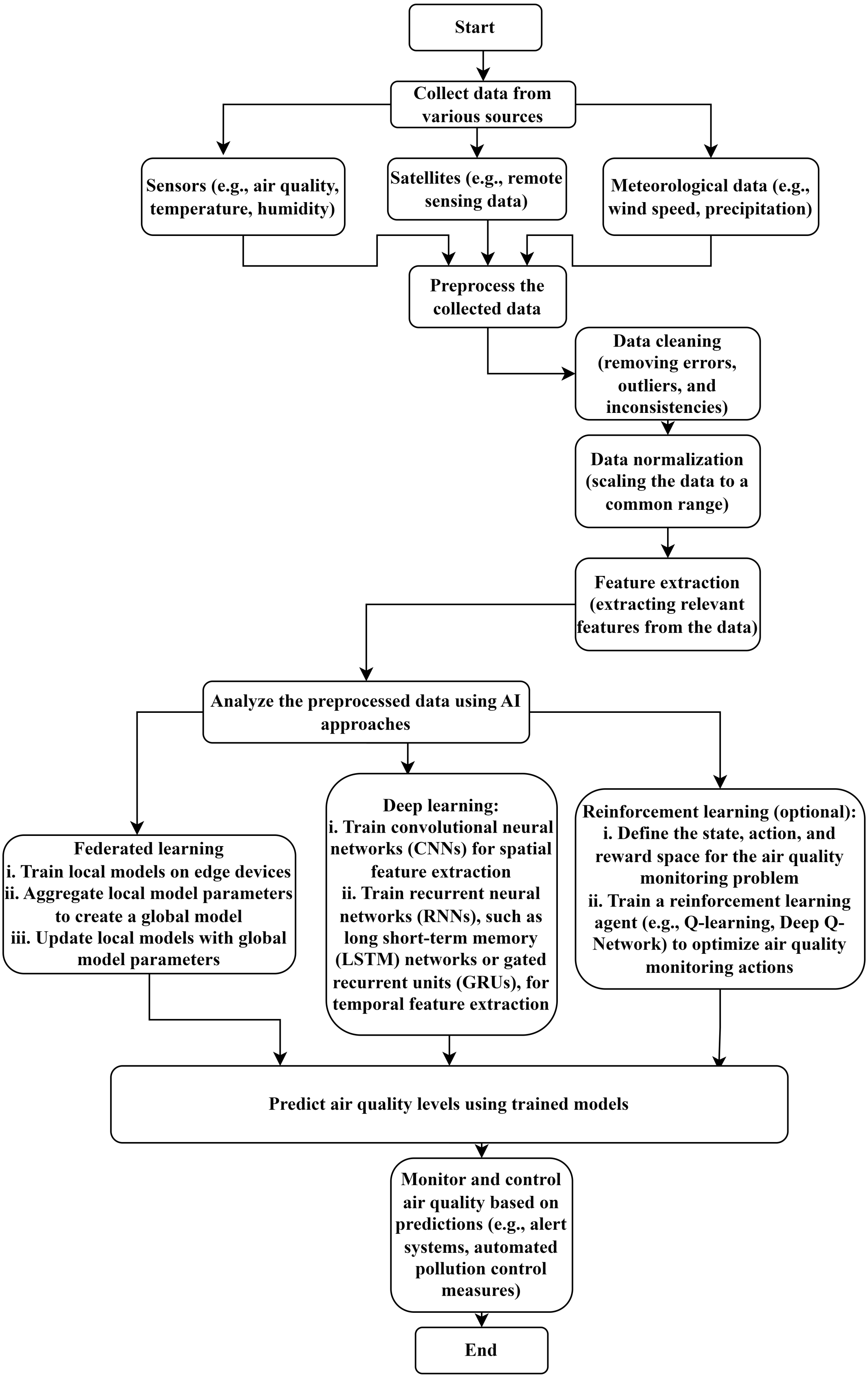


Figure 5. Flowchart for collecting, preprocessing, analyzing, and predicting air quality data using edge computing and AI approaches.

Following the data collection phase, the system dives into a rigorous preprocessing routine where data cleaning operations are conducted to ensure the removal of any errors, outliers, or inconsistencies that might compromise the accuracy of the predictions. Furthermore, data normalization procedures are carried out to scale the data to a common range, a critical step for the efficient training of machine learning models. This phase also includes a feature extraction process, where the system identifies and extracts the most salient features from the data set, a move that enhances the efficacy of the models developed later. Depending on the nature of the data (temporal or spatial), various types of features can be extracted. For example, for time-series data, it is common to use lag variables and rolling mean statistics. With deep learning, convolutional neural networks and auto-encoders can be trained to extract low-dimensional representations and feature maps.

Once the data have been adequately preprocessed, the system progresses to employ advanced AI techniques for a thorough analysis. This includes the use of federated learning, which promotes data privacy and the utilization of computational resources at the edge of the network by training local models on edge devices. These local model parameters are then aggregated to construct a global model, and the local models are updated with these global parameters. Deep learning techniques are also applied, with convolutional neural networks (CNNs) being employed for spatial feature extraction and recurrent neural networks (RNNs), like long short-term memory (LSTM) networks or gated recurrent units (GRUs), for temporal feature extraction. Optionally, reinforcement learning might be used to train a model that learns to interact with the environment and make decisions to maximize a reward. Following this, the performance of the various models is evaluated. Based on this evaluation, the system proceeds to make critical decisions. These can include predicting future environmental conditions, issuing alerts for extreme weather events or high pollution levels, and providing essential insights that can inform and shape environmental policy.

3.1. Edge Computing Techniques for Air Quality Monitoring and Predictions

The integration of edge computing with intelligent methodologies has emerged as a transformative tool for optimizing air quality, enhancing efficiency, monitoring, prediction, and ensuring data security and privacy, as evidenced by the studies outlined in the following sections.

Table 3 summarizes the diverse methodologies in the rapidly evolving field of edge computing for air quality monitoring. Methods range from container-based virtualization and federated compressed learning to advanced machine learning algorithms tailored for edge devices (e.g., [53,54,62,63]). These methodologies have shown promising applications in various domains, including general environmental monitoring, smart cities, indoor environments, and even niche sectors like the construction industry (e.g., [54,64–66]). Notably, these edge computing solutions offer substantial benefits, such as real-time data processing, high-resolution analysis, reduced power consumption, and enhanced data security (e.g., [53,54,62,67]). Future research should focus on developing adaptive and decentralized algorithms that can function efficiently on low-cost edge devices, particularly in resource-constrained settings. Additionally, the incorporation of game theory and other mechanisms for trust management will be critical for ensuring the reliability and security of these systems (e.g., [68]). As air quality monitoring becomes increasingly complex and critical, edge computing and intelligent methodologies are poised for transformative growth, warranting more in-depth, multidisciplinary investigations.

Table 3. Comparative analysis of edge computing methods in air quality monitoring (AQM) studies.

Reference	Year	Application Domains	Methodologies Utilized/Proposed	Benefits in AQM
[53]	2021	Environmental monitoring at universities	Implements container-based virtualization via Docker to isolate application environments. Kubernetes is used for orchestrating these containers, employing a cache-centric design that utilizes an etcd database for maintaining the cluster state	• Utilizes container-based virtualization to isolate applications, enabling optimized resource utilization. Leverages Kubernetes' cache-centric design for responsive real-time data processing and distribution
[54]	2021	General environmental monitoring	Leverages 5G for low-latency, high-bandwidth data transmission. Utilizes machine learning models, specifically a support vector machine (SVM) for sensor calibration and a convolutional neural network (CNN) for hyperspectral image analysis	• Integrates 5G for ultra-fast data transmission coupled with two distinct AI systems for sensor calibration and hyperspectral image analysis, ensuring real-time, high-resolution, and accurate air quality analysis
[67]	2018	Domestic and office environments	Utilizes low-power IoT sensors coupled with a proprietary edge computing framework. Employs dynamic load-balancing algorithms and thread-pool management to optimize CPU usage with battery-powered sensors	• Applies IoT sensors and edge computing for real-time monitoring, achieving up to 80% data precision • Distributes computational tasks to edge nodes, reducing overall system power consumption by 23%
[64]	2019	Smart cities	Deploys Raspberry Pi devices with ARM CPUs running lightweight Linux distributions serving as edge nodes. Integrates IoT sensors through GPIO pins and utilizes MQTT protocol for data communication. Edge computing algorithms are employed for localized data analytics to reduce latency	• Utilizes cost-effective Raspberry Pi devices coupled with edge computing to minimize cloud data storage costs and latencies • Facilitates real-time air quality insights for smart city applications
[68]	2020	Smart cities, urban environments	Combines evolutionary game theory and Lyapunov theory for dynamic trust management. Utilizes Kalman filtering for noise reduction in data and employs a Markov decision process for generating dynamic black-and-white lists based on evolving trust scores	• Introduces a trust management system using evolutionary game theory and Lyapunov theory • Manages the selection of trustworthy IoT devices dynamically, enhancing system resilience against malicious attacks
[62]	2021	Smart cities, wireless sensing networks	Adopts Federated compressed learning (FCL), which combines compressed sensing with federated learning Employs orthogonal matching pursuit (OMP) for sparse signal reconstruction and differential privacy techniques to add noise to the federated model, ensuring privacy	• Combines compressed sensing techniques with federated learning under the FCL framework • Achieves over 95% reduction in data consumption while ensuring privacy. Optimized for large-scale wireless sensor networks
[63]	2021	Remote areas, edge devices	Utilizes nonlinear autoregression with exogenous input (NARX) as a foundation. Enhances it by incorporating long short-term memory (LSTM) networks for capturing long-term dependencies and extreme gradient boosted random forests (XGBRFs) for an ensemble approach to increase robustness and reduce variance	• Offers rapid response to air quality changes and balances between computational speed and predictive accuracy • Enables real-time and comprehensive monitoring and prediction of air quality
[65]	2022	Construction industry	Integrates edge computing nodes equipped with specialized ASICs for real-time data processing and building information modeling (BIM) for 3D environmental modeling. Utilizes Bayesian networks for prediction and applies decision trees for generating real-time alerts and responses based on sensor inputs	• Holistic approach to air pollutant monitoring, prediction, and management at construction sites • Enhances predictive analytics, alerting, and emergency response mechanisms
[66]	2023	Indoor environments	Utilizes bi-directional long short-term memory (LSTM) networks to model temporal dependencies in IAQ metrics. Incorporates faster R-CNN for real-time object detection to identify and monitor potential pollutant sources, feeding these data back into the LSTM for dynamic modeling	• Real-time prediction of indoor air quality • Adaptability across various spaces and improves energy efficiency

3.1.1. Optimization and Efficiency

Centering on the optimization and efficiency of air pollution management, the authors in [53] employed container-based virtualization for achieving optimal efficiency and real-time data processing at Tunghai University. Simultaneously, in [64], the authors demonstrated a proficient real-time air pollution detection system using affordable Raspberry Pi devices in Tacoma, highlighting both its cost-effectiveness and potential in smart city applications. In this context, the cost-effective solution presented in [67] could prove valuable for domestic offices and industrial settings, enhancing data precision by up to 80% and reducing power consumption by 23%.

3.1.2. Real-Time Monitoring and Predictive Solutions

The role of real-time monitoring and predictive systems is pivotal in detecting air pollution and hazards. In this regard, the authors of [54] emphasized the fusion of 5G, edge computing, and AI to deliver real-time, high-resolution air quality analysis, highlighting the significance of prompt interventions for public health. Additionally, the authors of [63] designed an IoT-enabled system tailored for edge devices and cloud environments that swiftly responds to shifts in air quality through a hybrid predictive model. Similarly, a unified platform integrating edge computing and building information modeling (BIM) technology was introduced in [65], streamlining pollutant monitoring in the construction industry while enhancing predictive and emergency response capabilities. Furthermore, the study in [66] proposed an intelligent edge computing-based indoor system that accurately forecasts and monitors key IAQ factors, underlining its adaptable potential across diverse spaces.

3.1.3. Security and Privacy

When contemplating data processing encompassing numerous interconnected IoT devices, data security and privacy persist as major challenges that necessitate attention while conducting environmental air quality monitoring. Consequently, the work in [68] introduced a trust management mechanism for IoT edge computing, enhancing collaboration and providing robust defense against threats. Similarly, an FCL framework for PM2.5 air quality monitoring was introduced in [62], placing emphasis on data sparsity and privacy concerns and remarkably diminishing data consumption by over 95%.

3.2. Multi-Access Edge Computing in Air Quality Monitoring

Multi-access edge computing (MEC) signifies a pivotal paradigm shift in data processing, extending cloud computing capabilities to the edge of the network. This approach brings computational resources closer to end users and data sources, providing edge intelligence and significant benefits for air quality monitoring, which requires real-time processing and analysis of extensive data from environmental and air quality monitoring sensors [49,69–74].

One key advantage of MEC is its capacity to handle data locally, mitigating latency and network congestion. This aspect is critical in urban environments where air quality can fluctuate significantly over short distances [50]. The European Union-funded "CityAir" project exemplifies this benefit, utilizing MEC to develop a mobile air quality monitoring system capable of real-time pollutant detection and analysis, thus offering precise, localized air quality data to city residents and decision makers. It offers value in terms of optimizing the deployment of environmental sensors and streamlining the energy consumption of the monitoring systems [75,76]. MEC also shines in its ability for distributed data handling and collaborative processing. These features are crucial in large-scale air quality monitoring projects involving multiple stakeholders [76,77]. An illustrative example is the "SmartAQnet" project in the UK, which leverages MEC to construct a distributed air quality monitoring system. This system promptly delivers air quality data to residents, businesses, and local authorities, exemplifying MEC's potential in fostering scalable, flexible, and interoperable data sharing and collaboration [78]. Despite the remarkable potential of MEC, its implementation also presents certain challenges. A notable concern is data privacy and security, given the sensitivity of air quality data, which often include geolocation and other potentially identifying information. It is therefore essential for MEC systems to incorporate robust data encryption and access control mechanisms [50].

Table 4 furnishes a detailed overview of MEC strategies alongside their respective applications within the domain of air quality monitoring. A prominent observation is the consistent focus on harnessing MEC for real-time data processing, underscoring its growing importance in facilitating prompt air quality assessments. The studies underscore MEC's advantages in mitigating latency, ensuring robust data security, and facilitating scalable data sharing, particularly within urban contexts. Nonetheless, challenges like infrastructure expenses, data privacy apprehensions, and the necessity for standardized orchestration architectures persist. Addressing these challenges while optimizing MEC's real-time capabilities is crucial, especially in expansive deployments like 5G networks and smart cities, offering substantial potential for advanced researchers. Avenues such as multi-task learning and open-source edge computing methods could serve as promising directions for future innovation and research.

In conclusion, MEC offers promising opportunities for enhancing the efficiency and precision of air quality monitoring by facilitating faster, more accurate, and localized data processing and dissemination. However, to realize its full potential, addressing key challenges, including data privacy and security, is of paramount importance. As the field evolves, further research and technological innovation are necessary to optimize MEC-based air quality monitoring systems and ensure they meet the requisite privacy and security standards for their widespread adoption [77,79,80].

Table 4. Summary and comparison of multi-access edge computing (MEC) strategies and applications in air quality monitoring.

Reference	Year	Methodology	Applications	Main Methods	Key Findings
[70]	2020	Use of MEC for real-time processing and analysis of data from air quality sensors	Green 5G networks for evolving air pollution monitoring systems	Promotes rapid data processing, reducing delays in air quality assessment and reporting	Implementing this system across large-scale 5G networks may face issues relating to infrastructure and cost
[49]	2021	Implementation of MEC capabilities to process data at the edge of the network	Open-source approaches to intelligent edge computing, used in diverse applications	Enhanced proximity to end users and data sources, resulting in superior service delivery in data-intensive applications	Full potential of the open-source approach to edge computing still needs exploration
[50]	2021	Application of MEC for local data handling, reducing latency and network congestion	Review of MEC architectures with a focus on data security and privacy	Efficient data management in congested networks, vital for urban air quality monitoring	Concerns about data privacy and security due to the sensitivity of air quality data
[81]	2021	Leverage MEC for computational offloading and multi-task learning	CityAir project in the European Union, a mobile air quality monitoring system	Real-time pollutant detection and analysis, precise localized air quality data	Integration of multi-task learning within MEC systems needs further research for optimization
[77]	2022	Utilization of MEC's distributed data handling and collaborative processing capabilities	Review of Kubernetes-based orchestration architectures for smart cities	Enables large-scale projects involving multiple stakeholders, facilitates scalable and flexible data sharing	Development of standardized orchestration architecture for edge computing remains a challenge

3.3. Federated Learning in Air Quality Monitoring

Federated Learning (FL) has begun to reshape the landscape of decentralized machine learning, demonstrating applicability to air quality monitoring [82,83]. The profound potential of this approach stems from its capacity to draw upon distributed datasets while preserving data privacy and optimizing computational efficiency. Developments in the federated learning space have been marked by various distinct approaches. Figure 6 graphically outlines the federated learning approach. This visual representation shows a central cloud server around which numerous local devices, such as IoT sensors and smartphones, are spread out. Arrows depict the flow of information from the cloud server to the local devices and vice versa. The server maintains a global model and at each round shares the model with candidate clients, which in turn use their local data to update the model. The model's updates from different clients are sent back to the server to be aggregated and incorporated into the global model [70,83].

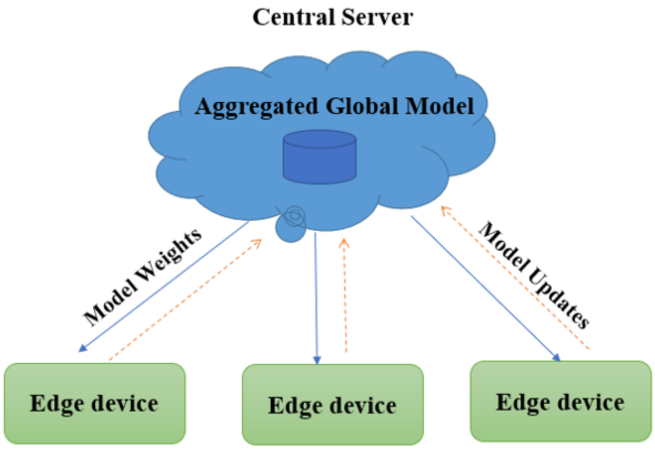


Figure 6. Federated learning approach.

Table 5 presents a summary of various federated learning frameworks that have been applied in air quality monitoring. Each row represents a distinct study, highlighting the authors, the specific federated learning framework deployed, the targeted applications, the primary methods or approaches employed, and the key findings. From advanced deep learning models, like CNN and LSTM [43,84], to unique client selection strategies [85], the table shows the breadth and diversity of techniques used within the context of federated learning. It also emphasizes the impact and benefits of these methods, such as significant improvements in computational efficiency and accuracy, superior performance compared to local training, privacy preservation, and secure and efficient data collection. The table serves as a quick reference for readers seeking to understand the current landscape of federated learning applications in air quality monitoring.

All these studies underscore the budding potential of federated learning frameworks in air quality monitoring. They offer promising pathways to addressing data privacy, transmission efficiency, and cooperative learning among distributed sensors. As we look ahead, the ongoing research in this area holds the promise of further augmenting federated learning capabilities, thereby unlocking new possibilities in environmental monitoring and beyond. Nguyen et al. [34], for instance, developed a framework that hinges on spatial averaging aggregation. Applied specifically to air pollution prediction, this framework leverages environmental sensor data and employs convolutional recurrent neural network (CRNN) models at the server side.

Liu et al. [86] brought forth another approach, creating a lightweight federated learning-based solution for aerial-ground air quality sensing. This method fuses multiple machine learning techniques, including convolutional neural networks (CNNs), graph convolutional networks (GCNs), long short-term memory (LSTM) networks, and fully connected (FC) layers. The innovation offers scalable, cost-effective, and fine-grained air quality monitoring, utilizing a swarm of unmanned aerial vehicles (UAVs). Another remarkable breakthrough came from Putra et al. [62], who presented federated compressed learning (FCL), a unique edge computing framework. Designed for PM2.5 predictions in smart city applications, this solution emphasizes security and efficiency in data collection by integrating compressed sensing (CS) with federated learning (FL).

In a more recent development, Xiang et al. [87] proposed S-KGCN, an edge-driven online IoT data search framework that employs knowledge graph convolutional networks for edge learning scenarios, including federated learning paradigms. This initiative highlights the versatility of federated learning in adapting to diverse application domains. An urban air quality monitoring dataset is used for evaluating the performance of the proposed method. Another federated learning approach is presented in [83] for imputing missing air quality data based on conditional generative adversarial networks (CGANs).

Table 5. Summary of federated learning frameworks applied in air quality monitoring.

Reference	Year	FL Framework	Applications	Main Methods	Key Findings
[43]	2018	Federated region-learning	PM2.5 monitoring in Beijing	Deep learning models like CNN and LSTM	Threefold improvement in computational efficiency, 5% increase in accuracy
[83]	2021	Federated GAN method	Intelligent air quality monitoring networks	Conditional GAN imputation method	Federated GAN method shows superior performance compared to local training
[34]	2021	Spatial averaging aggregation FL approach	Air pollution prediction model training with sensor data	Federated learning with CRNN	Proves that new participating areas can train and quickly obtain optimized local models using transferred common global model
[86]	2021	Lightweight federated learning-based framework	Privacy-protected AQI monitoring mobile application	Federated learning with a combination of CNNs, GCNs, LSTM networks, and FC layers	Proposes a framework for fine-grained air quality monitoring using UAVs
[62]	2021	Federated compressed learning (FCL)	PM2.5 air quality monitoring in smart cities	Federated learning with LSTM network	Demonstrates that the proposed FCL framework enables efficient data generation while ensuring data privacy for PM2.5 predictions
[47]	2021	Federated learning with autonomous UAVs	Air quality index (AQI) prediction in IoT environment	Hazardous zone detection, UAV-based data collection, and FL-based machine learning models	The proposed federated learning framework effectively identified hazardous zones and predicted AQI with higher accuracy and lower latency
[88]	2021	Federated learning with UAV swarm networks	Air quality index (AQI) prediction	UAV swarm-based data collection, FL-based machine learning models	Feasibility of using UAV swarms for AQI prediction with federated learning, which yielded promising results
[89]	2021	General federated learning	Estimation of air quality	Standard FL-based machine learning model	Successful application of FL for air quality estimation with high accuracy
[83]	2021	Federated conditional generative adversarial nets (FedCGANs)	Imputation of missing air quality data	FedCGAN-based imputation method	Effectively imputed missing air quality data, significantly improving the accuracy of air quality prediction
[85]	2022	Multi-model federated learning	Air quality index prediction	Ranklist-Multi-UCB and Pareto-Multi-UCB client selection strategies	Emphasizes the potential of multi-model federated learning
[90]	2022	Seminal stacked long short-term memory networks	PM forecasting	SS-LSTM networks for PM forecasting	SS-LSTM model outperformed all other models in PM forecasting
[91]	2022	General federated learning	Air quality index (AQI) prediction	Review and assessment of FL-based models for AQI prediction	Overview of FL techniques for AQI prediction, emphasizing the effectiveness of such methods
[59]	2022	Federated learning with BGRU model	Air quality prediction for smart cities	FL with BGRU model	Achieved significant results in predicting air quality for smart cities
[87]	2023	Edge-driven online IoT data search framework	Air quality monitoring in urban areas	Knowledge graph convolutional networks (KGCNs) for IoT data search	Proposes an edge-driven online IoT data search framework for edge learning scenarios, including federated learning paradigms
[82]	2023	Various FL frameworks	Smart city applications	Federated learning and machine learning algorithms	FL offers privacy preservation and sensitive information protection
[92]	2023	Secure federated learning with optimized LSTM	Cross-domain prediction of air pollutant concentration	Secure federated learning approach combined with optimized LSTM	Effective cross-domain prediction of air pollutant concentrations

3.4. Integration of FL and MEC in Air Quality Monitoring

The decentralized and collaborative advantages of federated learning (FL) and multi-access edge computing (MEC) mark the advent of a new era in data processing and computation, with the potential to revolutionize air quality monitoring systems [49,70,91,93]. MEC decentralizes computation and storage, effectively bringing these capabilities closer to the data source. Combined with FL's capacity to train machine learning models across a multitude of decentralized devices, this forms a powerful combination that enhances the effectiveness and precision of air quality monitoring systems [75,94].

The fusion of FL and MEC offers several benefits. Primarily, it allows efficient handling of large-scale, geographically dispersed data. MEC's ability to perform localized data processing alleviates latency and congestion issues, while FL's decentralized learning enables model training on edge devices, thereby optimizing predictive capabilities through the utilization of local data diversity [79,95]. This synergy between FL and MEC not only boosts computational efficiency but also reduces data transmission requirements, thereby diminishing network loads. This is especially critical in air quality monitoring systems, where devices are scattered across extensive geographic areas and data transmission can become a bottleneck [70,78]. Additionally, this integrated approach offers enhanced privacy preservation and data security. During the learning process, sensitive air quality data remain on local devices, bolstering privacy and security [50,96]. Emerging practical applications of this integrated FL–MEC approach include the utilization of an edge-native intelligence system for large-scale urban air quality monitoring networks, which provides real-time, hyper-localized data [35,55].

Despite these benefits, the integration of FL and MEC introduces certain challenges. For instance, synchronizing FL processes across various edge devices while maintaining computational efficiency can be complex. Furthermore, despite the inherent data privacy benefits of FL, complete data security across a distributed network remains a concern, particularly for air quality data containing sensitive location-specific information [50,97]. In summary, the integration of FL and MEC shows significant promise for enhancing air quality monitoring systems. By capitalizing on the strengths of both technologies, we can develop robust, efficient, and secure solutions capable of processing and analyzing large and diverse datasets in real time. However, addressing these inherent challenges is paramount for the successful implementation of these innovative approaches. Continuous research is essential to refine these integrated systems and establish standardized practices for their optimal use [43,90,98].

Table 6 encapsulates pivotal findings from a range of studies [34,35,50,55,62,70,78,79,82,83,91,95–97,99] that underscore the essential integration of FL and multi-access MEC within the domain of air quality monitoring for smart cities. Thorough examination of the presented literature clearly reveals that the integration of FL and MEC effectively addresses crucial research gaps, especially in terms of enhancing data privacy, efficiency, and security, as highlighted in [70,82,97]. This integration holds the potential to revolutionize data acquisition methods [62,95] while simultaneously establishing a scalable and sustainable model [78,79], rendering it particularly pertinent for future research directions. Given the trajectory of smart city applications and their mounting complexities, forthcoming research should delve deeper into optimizing this integration [62,96], creating an environment where data consumption is minimized, privacy is prioritized, and data quality is elevated [99]. The synthesis outlined in Table 6 can thus stand as a cornerstone for scholars, compelling them to build upon these foundational insights and further push the boundaries in the ever-evolving field of smart city air quality monitoring.

Table 6. Main features of integrating FL and MEC in air quality monitoring.

Feature	Description	References
Scalability and efficient data handling	The integrated FL–MEC approach efficiently handles vast, geographically diverse data, mitigating latency and congestion issues	[62,79,95]
Reduced network load	By minimizing the need for extensive data transmission and enabling localized data processing, FL and MEC reduce network loads, which is beneficial for wide-area air quality monitoring systems	[70,78,82]
Privacy and data decurity	FL keeps sensitive air quality data on local devices during the learning process, preserving data privacy. Moreover, MEC's local processing capability reduces sensitive data transmission, ensuring security	[50,62,96]
Real-time hyper-localized data	FL–MEC integration provides real-time, hyper-localized data for large-scale urban air quality monitoring networks, improving predictive capabilities	[35,55,91]
Challenges	Reducing operational complexity, synchronizing FL processes across multiple edge devices, handling stragglers, and maintaining computational efficiency are inherent in the FL–MEC integration. Despite FL's data privacy benefits, ensuring complete data security across a distributed network remains a concern	[34,50,83,97,99]

3.5. Model Construction Paradigms

3.5.1. Machine Learning for Air Quality Monitoring and Forecasting

Artificial intelligence (AI) strategies have emerged as significant enablers in improving the accuracy and trustworthiness of air quality monitoring and forecasting, capitalizing on their capacity to handle large datasets and extract underlying pattern. Machine learning (ML), a core AI discipline, employs a spectrum of statistical algorithms that provide computer systems with the ability to autonomously learn from data and progressively enhance their performance in tasks [4,100]. This iterative learning capability without explicit programming has positioned ML as a pivotal tool in air quality monitoring and forecasting, delivering innovative methodologies to analyze extensive datasets, extract complex patterns, and predict future air quality trends [27,101].

In the field of air quality monitoring, ML finds its significance in developing predictive models. These models leverage historical air quality data to predict future air quality indices, a process marked by significant complexity due to the involvement of numerous factors, including meteorological conditions, traffic flow, and emission data [84,102,103]. The integration of these variable datasets through ML algorithms enhances the predictive capability, offering improved accuracy and timeliness in air quality forecasting [104]. Furthermore, researchers have employed ML techniques for source apportionment studies, an essential aspect of air quality monitoring. By analyzing real-time air quality data, ML algorithms can determine potential emission sources, distinguishing between natural and anthropogenic contributors [66,105]. This identification is crucial for targeted pollution mitigation strategies, significantly improving the effectiveness of air quality management [102,103].

Various ML techniques, such as regression models, decision trees, and support vector machines, have been used in the field, each presenting unique strengths. For example, linear regression models, owing to their simplicity and interpretability, have been extensively used for continuous data predictions, such as forecasting pollutant concentrations [106]. On the other hand, decision trees and support vector machines have shown exceptional performance in classification tasks, such as identifying pollutant types [107,108]. The effectiveness of these ML techniques in air quality monitoring, however, is highly dependent on the quality and quantity of the input data. Data imbalances and irregularities can significantly impact model performance, leading to inaccurate forecasts or misclassifications [102]. To counter these issues, various preprocessing techniques have been utilized, such as data augmentation, normalization, and outlier detection, to ensure data integrity and enhance model performance [109,110].

Table 7 delves into the transformative role of ML in both air quality monitoring and forecasting. Leveraging data from Web-based platforms, such as Google Traffic, has demonstrated a remarkable ability to significantly enhance predictions of urban air pollution, consistently outperforming conventional techniques by margins of 6.5% to 8.5% [107,111]. The integration of ML and affordable sensors notably enhances air quality monitoring, especially in settings constrained by limited resources, even though regular calibration remains necessary [112,113]. Specific ML models, such as random forest and Gaussian process regression, have garnered attention for their proficiency in predicting air quality across diverse urban environments [114–116]. Recognizing the profound impact of human-induced factors, including traffic and urbanization, holds paramount importance in devising impactful interventions [117]. Particularly noteworthy is the synergy between ML and cost-effective technologies, which has emerged as a valuable strategy, particularly beneficial for economically challenged and developing regions, to effectively address urban air quality challenges [107,111,113,115].

Table 7. Comparative summary of traditional ML techniques used in air quality monitoring and forecasting.

Technique	Applications	Strengths	Limitations	References
Linear regression models	Used for continuous data predictions, such as forecasting pollutant concentrations	Simple, interpretable, and handle continuous data well	Assume a linear relationship between variables, may not capture complex patterns	[111–115]
Decision trees	Used for classification tasks, such as identifying pollutant types	Handle categorical data well, easy to interpret	Prone to overfitting, may not handle continuous data well	[107,108]
Support vector machines (SVMs)	Used for both regression and classification tasks in pollutant concentration prediction and pollutant type identification.	Can handle large feature spaces, effective in high-dimensional spaces	Computationally expensive, require careful parameter tuning	[116]
Random forests	Used for both regression and classification tasks, highly applicable for source apportionment	Handle large datasets well, reduce overfitting compared to decision trees	Computationally intensive, less interpretable than decision trees	[117]
Neural networks	Used for predicting future air quality levels and recognizing complex patterns	Can model nonlinear relationships and handle large datasets well	Require large amounts of data, prone to overfitting	[118]

3.5.2. Deep Learning for Air Quality Monitoring and Forecasting

Deep learning (DL) is a branch of machine learning demonstrating an advanced layer of complexity compared to traditional machine learning. One particular merit of deep learning approaches is their capability for hierarchical feature extraction in several cascaded layers. Initial layers are more related to general low-level features that are common across multiple similar tasks and transfer learning layers towards the end can be tuned for various tasks. DL has been gradually surfacing as a potent tool for air quality monitoring and forecasting. Its inherent ability to autonomously identify intricate patterns and relationships within large-scale datasets, an attribute essential to process the enormous volume and complexity of environmental data, sets it apart from other techniques. In the arena of air quality prediction and monitoring, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are the two prominent deep learning models utilized, each showcasing distinct advantages.

Table 8 offers an analytical overview of diverse deep learning techniques harnessed for air quality forecasting. CNNs have demonstrated remarkable performance in predicting $PM_{2.5}$ concentrations and AQI by utilizing satellite imagery and ground sensor data, underlining their prowess in handling geo-spatial datasets [81,119]. LSTM networks, a subset of RNNs, also exhibit strong forecasting capabilities, particularly for $PM_{2.5}$ concentrations, AQI, and overall air pollution levels, often surpassing traditional time-series models [103,120,121]. Concurrently, ensemble and multi-model deep learning approaches present a significant enhancement in prediction accuracy [85,122,123]. For researchers venturing into this domain, it is imperative to explore CNN adaptability across varying data scales; delve deeper into LSTM's robustness, feature selection, and spatiotemporal adaptability; and rigorously investigate ensemble methodologies while concurrently addressing model interpretability.

Table 8. Summary of various deep learning techniques.

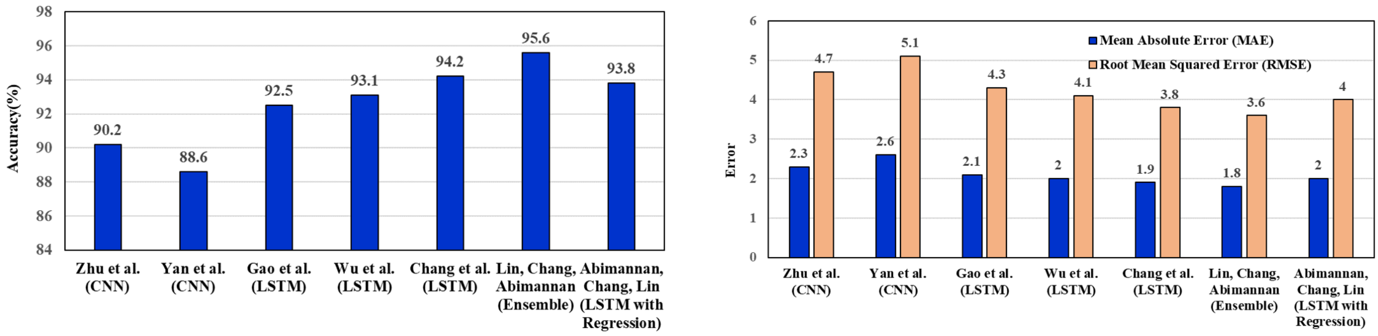
Technique	References	Dataset Utilized	Applications	Key Finding	Considerations
Convolutional neural networks (CNNs)	[81,119]	Satellite imagery and meteorological data; ground-based sensor data	PM2.5 concentration prediction; AQI prediction	Demonstrated superior performance compared to traditional ML techniques, providing accurate PM2.5 predictions and AQI predictions	Studies emphasize the effectiveness of CNNs in leveraging both geospatial and sensor data, suggesting further exploration and study of the adaptability and scalability of CNNs
Recurrent neural networks (RNNs), long short-term memory (LSTM) networks	[103,120,121]	Air quality data	PM2.5 concentration forecasting; AQI forecasting; air pollution forecasting	LSTM models showed remarkable performance in forecasting PM2.5 concentrations, AQI, and air pollution, outperforming traditional time-series models and even demonstrating enhanced results with the integration of multivariate regression.	Studies underline the potential of LSTM networks for air quality forecasting but highlight the need for further research on LSTM model robustness, the effect of feature selection, and adaptability to diverse spatiotemporal resolutions
Ensemble and multi-model deep Learning	[85,122,123]	Air quality data	Air quality forecasting	Introduction of an ensemble of multifeatured and multi-model deep learning models substantially enhanced prediction accuracy	Studies showcase the promise of ensemble and multi-model methodologies but call for further comprehensive studies on ensemble methodologies and their impact on model interpretability

CNNs have etched a substantial mark in air quality monitoring through the analysis of diverse data sources, like satellite imagery and ground-based sensor data. Yan et al. [81] employed CNNs to effectively handle various data types (satellite imagery and meteorological and sensor data), demonstrating the models' prowess in extracting complex spatial features. Another study by Suriya et al. [124] leveraged a CNN-based model to predict PM2.5 concentrations using satellite imagery coupled with meteorological data, outshining traditional machine learning models in terms of prediction accuracy.

Conversely, RNNs—specifically, long short-term memory (LSTM) networks—have found applications in air quality forecasting. Their ability to model temporal dependencies in air quality data makes them a fitting choice for such tasks. LSTM's capability to model temporal dependencies is highlighted in a number of studies [103,120,121,123], and these models outperformed traditional time-series models in air quality forecasting accuracy. For instance, Gao et al. [120] implemented an LSTM-based model for forecasting PM2.5 concentrations. The model outclassed traditional time-series models, exhibiting higher accuracy. Similarly, Wu et al. [103] also developed an LSTM-based model, achieving remarkable accuracy in AQI forecasting. Dey et al. [59] proposed a decentralized FL-based system utilizing the bidirectional gated recurrent unit (BGRU). The integration of multivariate regression in air pollution forecasting, as demonstrated in [121], is another vital milestone, reinforcing the prowess of LSTM networks in this field. Adding to the repository of significant studies, Chang et al. [123] devised a hybrid stacking ensemble model for air pollution forecasting of PM2.5 and PM10, exhibiting high levels of precision. Lastly, the work in [122] introduced another ensemble of multifeatured deep learning models for air quality forecasting, thereby enhancing prediction accuracy and stability and underscoring the potential of integrating different models for enhanced outcomes.

Figure 7 provides insights into the efficacy of various deep learning techniques in air quality forecasting gleaned from an examination of performance metrics across different studies. It is particularly notable that LSTM-based models, including those implemented in [103,120,123], demonstrate consistently high accuracy, ranging between 92.5% and 94.2%. This accuracy increases further to 93.8% when the LSTM network is integrated with multivariate regression, as evidenced by the study conducted in [121]. The ensemble model approach, developed by Lin, Chang, and Abimannan [123], outperforms individual deep learning models, reaching an impressive peak accuracy of 95.6%. This indicates the potential of leveraging the strengths of multiple models to enhance predictive performance. Further examination of error metrics reveals that LSTM-based models consistently maintain a lower error rate, exhibiting minimal mean absolute error (MAE) and root mean squared error (RMSE) values, suggesting a closer fit to the actual data. The ensemble model approach continues to showcase its superior performance, yielding the smallest MAE and RMSE values. Consequently, these insights not only reinforce the importance of using advanced deep learning techniques in air quality monitoring and forecasting applications but also underscore the potential benefits of technique fusion and ensemble approaches in achieving superior prediction outcomes.

In conclusion, deep learning techniques hold vast potential to revolutionize the domain of air quality monitoring and forecasting. However, due to the inherent complexities and diversities of environmental data, additional research efforts are necessary to harness the full potential of these advanced techniques.



(**a**) Accuracy (**b**) MAE and RMSE

Figure 7. Outcomes of respective studies comparing the performance of various deep learning techniques in air quality forecasting [81,103,119–123].

4. Challenges and Requirements for Developing and Implementing Real-Time Air Quality Monitoring Systems

While edge computing and intelligence offer many benefits for real-time air quality monitoring, they also have some limitations, such as limited processing and storage capacity for complex algorithms [125], limited scalability [126], constrained connectivity in remote areas [127], limited support for high-volume and heterogeneous data streams [128], and restricted interoperability with existing systems [129,130]. Table 9 summarizes various challenges facing the implementation of a real-time air quality monitoring system. As discussed in the following sections, these challenges include ensuring the accuracy, completeness, and timeliness of data from diverse sources; achieving spatial and temporal coverage across urban areas; and enhancing the accuracy and interpretability of air quality prediction models. Additionally, resource-constrained edge computing, efficient communication and networking, federated learning for collaborative data processing, and robust data security and privacy measures are essential. Designing scalable and flexible architectures, fostering interdisciplinary collaboration, and involving stakeholders in system development and evaluation are also crucial for successful implementation in research-focused air quality monitoring initiatives.

Table 9. Summary of challenges and requirements for developing and implementing real-time air quality monitoring systems.

Challenge/Requirement	Description	Example Methods
Data quality and availability	Ensuring accuracy, completeness, and timeliness of air quality data collected from various sources	Outlier detection [131], data imputation [132], data integration and fusion [133]
Spatial and temporal coverage	Providing comprehensive coverage of air quality parameters across urban areas and at different spatial and temporal resolutions	Spatiotemporal interpolation [104], time series analysis [133], spatial data analysis [134]
Modeling and prediction accuracy	Developing accurate, reliable, and interpretable air quality prediction models that can adapt to varying conditions	Deep learning models (e.g., GRU, LSTM, CNN) [59], ensemble methods [135], hybrid and decomposable models [136]
Edge computing and resource constraints	Efficiently processing and analyzing air quality data on resource-constrained edge devices while reducing latency and resource utilization	Edge AI [137], computation offloading [130], edge-based data preprocessing [134]
Communication and networking	Ensuring reliable, efficient, and low-latency communication among sensors, devices, and data centers	Low-latency communication protocols [138], IoT networks [134], cloud–edge orchestration [136]
Federated learning and collaboration	Enabling decentralized collaborative learning and data processing among different stakeholders without compromising data privacy	Federated learning [139], collaborative training [140], FL for the IoT [141]
Data security and privacy	Protecting the confidentiality, integrity, and availability of air quality data and preserving user privacy	Differential privacy [142], secure multi-party computation [143]
System scalability and flexibility	Designing scalable and flexible architectures that can accommodate growing data volumes, expanding urban areas, and changing environmental conditions	Scalable architectures [136], dynamic model optimization [141], cloud–edge IoT framework [136]
Interdisciplinary collaboration and stakeholder involvement	Fostering interdisciplinary research and involving key stakeholders in the design, implementation, and evaluation of air quality monitoring systems	Cross-domain collaboration [135], stakeholder engagement [141], FL in robotics [142]

4.1. Data Quality and Preprocessing

One of the primary challenges in developing real-time air quality monitoring systems is ensuring data quality and proper preprocessing. Data collected from sensors may demonstrate noise, missing values, or inaccuracies due to various factors, such as faulty sensors, environmental conditions, or communication issues [83,94]. These data issues need to be addressed before feeding them into prediction models. Data preprocessing techniques like data cleansing, imputation, and normalization can be employed to tackle these challenges [132]. Ensuring data quality begins with proper sensor calibration and maintenance. Regular calibration of sensors can minimize errors and improve data accuracy [135]. In addition, advanced sensor fusion techniques can be employed to combine data from multiple sensors to enhance the overall quality of the collected data [133,144–147]. The steps involved in sensor calibration includes initial calibration, regular calibration, and calibration verification to ensure accurate data collection. Proper sensor maintenance can help minimize errors due to factors such as environmental conditions, sensor drift, and aging, thereby improving data quality and reliability for real-time air quality monitoring.

Data cleansing is an essential step in preprocessing, where erroneous and inconsistent data points are identified and corrected or removed. Various methods, such as outlier detection, clustering, and rule-based techniques, can be used to detect and correct data anomalies [133]. However, the application of these methods is challenging and should be treated cautiously in order not to lose important information. For example, sensors may report spikes due to temporary events such as passing vehicles, but these spikes are important and should not be eliminated by an outlier detection scheme. Handling missing data is another crucial aspect of preprocessing. Imputation methods, such as mean imputation, regression imputation, and more advanced techniques, like matrix completion or generative adversarial networks, can be used to fill in missing values [83]. Normalization techniques, such as min–max scaling and Z-score normalization, can be applied to ensure that the data are on a consistent scale and prevent the domination of certain features during model training [132].

4.2. Model Training and Deployment

Developing accurate and reliable prediction models for air quality monitoring is a major challenge. While numerous machine learning algorithms and techniques have been proposed in the literature [59,104,131], model training and deployment in real-time systems can be computationally intensive, especially with deep learning models. To address this issue, federated learning and edge computing have been proposed as potential solutions [138,139,142].

Federated learning is a decentralized approach to training machine learning models that allows data to remain on local devices while model updates are shared across the network [140]. This approach can reduce communication overhead and the computational burden on central servers, making it suitable for large-scale, distributed air quality monitoring systems. Moreover, privacy-preserving federated learning techniques, such as homomorphic encryption, secure multi-party computation, and differential privacy, can be employed to protect sensitive data during the model training process [143].

In addition, edge computing is a paradigm that brings computation and data storage closer to the devices where data are generated, thereby reducing the need for data transmission to the cloud [7]. By offloading computation tasks to edge devices, such as IoT gateways or edge servers, real-time air quality monitoring systems can achieve lower latency and improved response times. Additionally, edge computing can help mitigate the impact of network congestion and increase the overall scalability of the system [37,130,148].

4.3. Scalability and Heterogeneity

Real-time air quality monitoring systems need to be scalable to accommodate the increasing number of IoT devices and sensors in urban environments [141]. Scalability is a major challenge, as it involves efficient management of resources, handling large volumes of data, and adapting to the changing dynamics of the system. Distributed architectures, such as edge computing, can help improve the scalability of air quality monitoring systems by distributing the computation workload among multiple edge devices [149]. Furthermore, federated learning can enhance scalability by enabling decentralized model training and reducing the need for centralized data storage [94].

Heterogeneity is another challenge in real-time air quality monitoring systems, as these systems often involve various types of sensors, devices, and communication protocols [135]. Ensuring interoperability among heterogeneous components can be a complex task. Standardization of communication protocols, data formats, and APIs can help address heterogeneity issues and facilitate seamless integration of different system components [137]. Moreover, employing middleware solutions that can abstract the underlying hardware and software heterogeneity can further improve the interoperability and flexibility of the system [136].

4.4. Data Security and Privacy

Data security and privacy pose significant challenges in real-time air quality monitoring systems due to the vast quantities of sensitive data they handle. For example, a single air quality monitoring system can process up to several terabytes of data daily, including potentially sensitive details, such as location information and personal data from users [150]. Ensuring the confidentiality, integrity, and availability of these data is paramount to guard against unauthorized access, tampering, and data loss. Implementation of robust security measures, like data encryption, access control mechanisms, and secure communication protocols, plays a critical role in preserving data security during both storage and transmission stages [150].

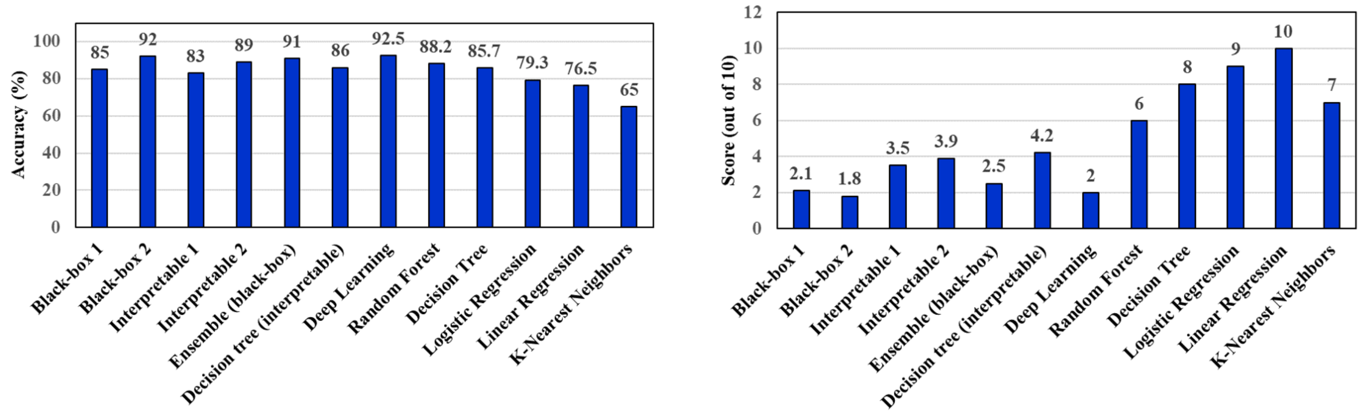
In addition to data security, maintaining user privacy is crucial, especially when applying federated learning approaches [140]. Privacy-preserving techniques, such as differential privacy, secure multi-party computation, and homomorphic encryption, can be employed to protect user data during the model training process [143]. These techniques help minimize privacy leaks during the exchange of model updates while still enabling collaborative learning among multiple devices.

4.5. Interpretable and Explainable AI Models

Interpretable and explainable AI models are crucial for real-time air quality monitoring systems, as they facilitate trust in the system by providing insights into the decision-making process and enabling users to understand the underlying reasons for the predictions [151]. Machine learning models used in air quality monitoring systems, such as deep learning and ensemble models, are often considered black-box models due to their complex nature and lack of transparency for the decision-making process [132]. Developing interpretable and explainable AI models can improve the transparency and accountability of the system, which is essential for regulatory compliance, user trust, and effective decision making.

Techniques such as local interpretable model-agnostic explanations (LIMEs) and Shapley additive explanations (SHAPs) can be employed to provide local explanations for individual predictions, while global explanation methods, such as feature importance ranking and partial dependence plots, can provide insights into the overall behavior of the model [152]. Furthermore, incorporating interpretable models, such as decision trees, linear regression, or sparse models, can enhance the explainability of the system without compromising prediction accuracy [133].

Figure 8 shows the performance of different types of machine learning models in real-time air quality monitoring systems. It highlights the trade-off between model accuracy and explainability and underscores the importance of choosing the appropriate type of model for real-time air quality monitoring systems based on their specific requirements and priorities. The black-box models, which include black-box 1, black-box 2, and the ensemble model, achieve high accuracy rates ranging from 91.0% to 92.5%. However, they score low in explainability, with scores ranging from 1.8 to 2.5 on a scale of 1–10. On the other hand, the interpretable models, which include interpretable 1, interpretable 2, decision trees (interpretable), and other linear models, score higher on explainability, ranging from 3.5 to 10, but achieve slightly lower accuracy rates ranging from 76.5% to 89.0%. The random forest model achieves the highest explainability score of 6 and an accuracy rate of 88.2%.



(**a**) Accuracy (**b**) Explainability score

Figure 8. Performance of interpretable and explainable models compared to black-box models in terms of accuracy and explainability. Data are based on a survey of recent studies [132,133,151].

4.6. Integration with IoT and Smart City Infrastructures

Integrating real-time air quality monitoring systems with the IoT in smart city infrastructures and public safety is essential for enabling seamless data collection, efficient resource management, and coordinated decision making across various urban domains [141,153–159]. To achieve this, it is necessary to design and implement interoperable and modular architectures that can facilitate integration with existing and future IoT and smart city solutions [7]. Middleware platforms and standardized APIs can help bridge the gap between heterogeneous system components, ensuring seamless data exchange and interaction among various IoT devices, sensors, and services [93,137,160,161,161]. Additionally, adopting a service-oriented architecture (SOA) can improve the modularity and flexibility of the system, enabling the integration of new services and applications as the smart city ecosystem evolves [138].

As depicted in Figure 9, various SOA characteristics, such as standardized interfaces, loose coupling, reusability, scalability, and interoperability, are critical for enhancing the modularity, flexibility, and integration of real-time air quality monitoring systems. The modularity scores range from seven to nine, with standardized interfaces receiving the highest score of 9, indicating that SOA enables greater modularity in real-time air quality monitoring systems. The flexibility scores range from six to nine, with loose coupling and interoperability receiving the highest scores of nine, indicating that these methods enhance the flexibility of the system. The integration scores range from seven to nine, with standardized interfaces and interoperability receiving the highest scores of nine, indicating that they facilitate better integration of the system with other smart city solutions. The scores for reusability and scalability are lower than the other SOA characteristics, indicating that they may have a lesser impact on the modularity, flexibility, and integration of the system. Overall, Figure 9 highlights the importance of adopting a service-oriented architecture in real-time air quality monitoring systems to improve their modularity, flexibility, and integration, ultimately leading to more efficient resource management and better decision making in smart city environments.

Based on Figure 10, it can be observed that the integration of advanced techniques and technologies, such as federated learning, edge computing, and privacy-preserving methods, can have a significant impact on the accuracy, reliability, and timeliness of air quality predictions in real-time air quality monitoring systems. Federated learning, which allows for model training with distributed data while preserving privacy, has the highest percentage increase in accuracy, with a 12.5% improvement over traditional machine learning models. Edge computing, which involves processing data closer to the source to reduce latency, has a moderate impact on accuracy, with a 5% improvement. Privacy-preserving methods, which protect sensitive data while allowing for analysis, have a lower impact on accuracy, with a 2% improvement. In terms of reliability, all three advanced techniques and technologies show an improvement, with privacy-preserving methods having the greatest impact with a 20% increase in reliability. Edge computing and federated learning also show improvements in reliability, with 10% and 5% increases, respectively. When it comes to timeliness, edge computing has the greatest impact with a 50% increase in speed. Federated learning and privacy-preserving methods also show improvements in timeliness, with 20% and 10% increases, respectively. Overall, these results highlight the potential benefits of integrating advanced techniques and technologies into real-time air quality monitoring systems. However, it is important to note that the specific impact may vary depending on the implementation and specific context.

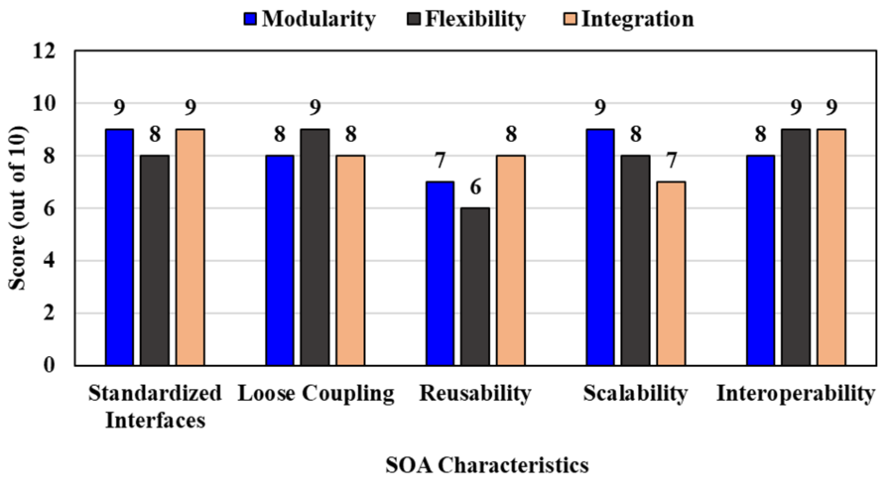


Figure 9. SOA characteristics of real-time air quality monitoring systems [7,137,138,141].

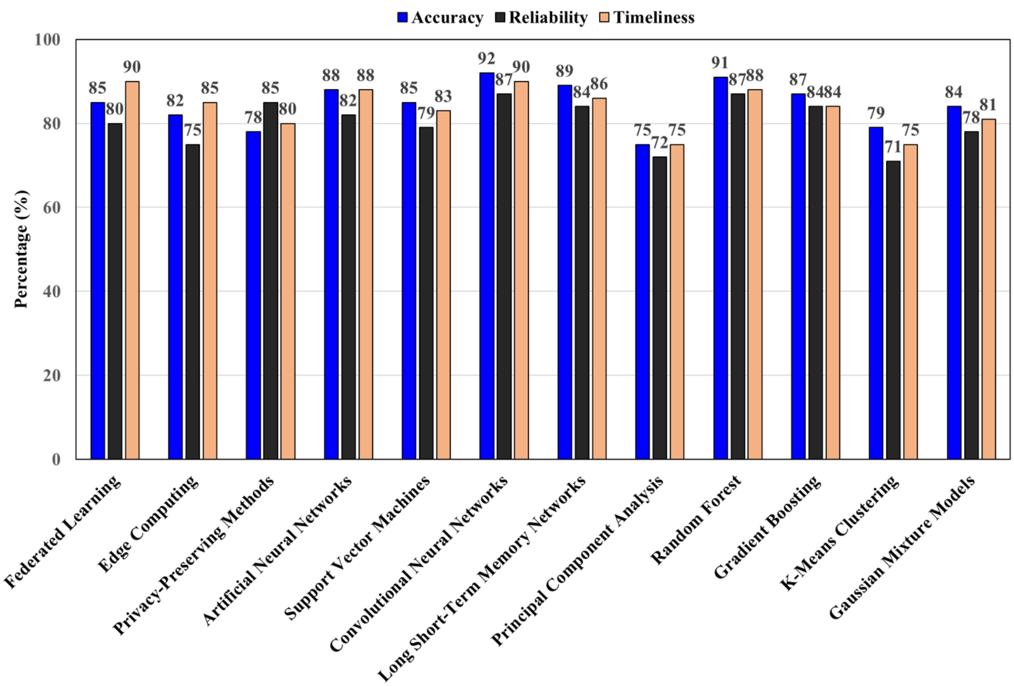


Figure 10. Impact of advanced techniques and technologies on real-time air quality monitoring systems [17,134,162–164].

4.7. Data Security and Privacy

Real-time air quality monitoring systems often rely on the collection, processing, and sharing of large amounts of data, which raises concerns about data security and privacy [142]. Ensuring the confidentiality, integrity, and availability of the collected data is crucial for maintaining user trust and compliance with data protection regulations [139]. Implementing robust security mechanisms, such as encryption, authentication, and access control, can help safeguard the system against unauthorized access, data breaches, and other security threats [143].

Moreover, preserving user privacy is essential when collecting and processing personal data, such as location information and user preferences [94]. Privacy-preserving techniques, such as anonymization, data obfuscation, and differential privacy, can be employed to protect user privacy while maintaining the utility of the data for air quality monitoring and prediction [149]. Furthermore, federated learning and secure multi-party computation methods can enable collaborative learning and data processing among different stakeholders without exposing sensitive information [59].

4.8. System Scalability and Flexibility

As cities continue to grow and evolve, air quality monitoring systems need to be scalable and flexible to accommodate increasing data volumes, expanding urban areas, and changing environmental conditions [140,165,166]. Designing modular and extensible architectures that can easily integrate new sensors, devices, and data sources is essential for ensuring the long-term adaptability of the system [136]. Furthermore, leveraging scalable data processing and storage techniques, such as distributed computing, data sharding, and data compression, can help accommodate the growing data demands and ensure the efficient operation of the system under varying loads [137]. In addition, adopting flexible modeling and prediction approaches that can adapt to new data and changing conditions can improve the system's ability to maintain accurate and reliable air quality forecasts as the urban environment evolves [131].

Figure 11 illustrates the relationship between system load and response time for a real-time air quality monitoring system with distributed computing capabilities. The x-axis indicates to the amount of data in GB loaded in the system for processing and analysis, which could represent high-resolution sensor data, an increased number of sensors contributing data, or longer periods of time over which data have been collected. The system employs a distributed architecture with a master node and several worker nodes to process and analyze large volumes of data in real time. A load balancing algorithm distributes the workload among the worker nodes based on their processing power and availability. Data sharding and compression techniques optimize data transfer and storage, ensuring efficient operation even under high loads. The data in the figure were collected using a benchmarking tool that simulated increasing data loads on the system and measured the corresponding response times. The results highlight the effectiveness of distributed computing techniques in maintaining efficient operation of real-time air quality monitoring systems under varying data loads.

A real-time air quality monitoring system with distributed computing capabilities can utilize a microservice architecture and Apache Kafka for message queuing, similar to the approach described in [136] for live video analytics. This system comprises multiple independent modules that can be deployed on separate nodes, enabling horizontal scaling as the data load increases. Each module runs in its container, providing isolation and enabling independent scaling.

The system uses load balancing techniques to distribute the incoming data stream across multiple nodes and minimize processing time [137]. The load balancer uses a weighted round-robin algorithm to assign the data streams to the available nodes based on their processing capacity. As the data load increases, the load balancer dynamically adjusts the weights to ensure that the processing load is evenly distributed across the available nodes. To ensure fault tolerance and high availability, the system may use Apache ZooKeeper for distributed coordination and leader election. ZooKeeper monitors node health and performs failover in case of a node failure or network partition. For data processing, the system utilizes Apache Spark for distributed data processing and machine learning. Spark uses a distributed computing model and in-memory processing to enable fast and efficient data processing even under high loads. The system employs various optimization techniques, such as caching, pipelining, and dynamic partitioning, to further enhance performance and reduce processing time. Overall, the system's distributed computing capabilities enable efficient processing of large volumes of data and maintain fast response times even under high loads, making it well suited for real-time air quality monitoring applications.

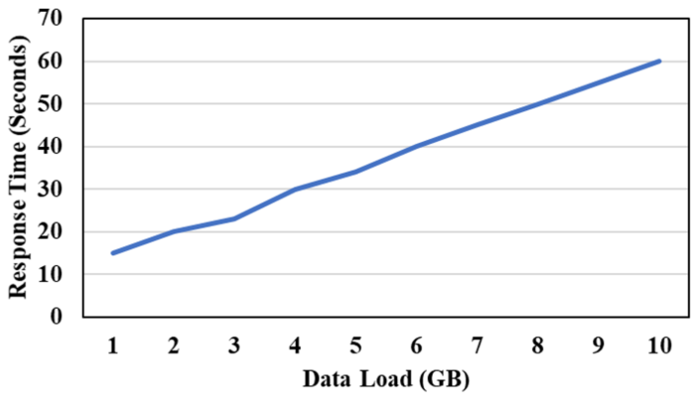


Figure 11. System load versus response time for real-time air quality monitoring system with distributed computing capabilities.

5. Insights for Future Research Directions

In recent years, federated learning (FL) has emerged as a promising solution for air quality monitoring and forecasting in smart cities. FL is a distributed machine learning paradigm that enables multiple parties to collaboratively train models with their local datasets without sharing raw data, thus preserving privacy and security. FL is especially suitable for air quality monitoring and forecasting because it can efficiently process large amounts of data generated by heterogeneous sources, such as sensors, weather stations, and social media, while preserving the privacy of individuals and organizations [138].

One of the future research directions in FL for air quality monitoring and forecasting is to develop more efficient and robust federated learning algorithms that can handle the heterogeneity and complexity of data in smart cities. Most FL algorithms assume homogeneous data sources and use simple aggregation methods, such as federated averaging (FedAvg), which can lead to biased or suboptimal models. To address this challenge, researchers can explore more advanced FL algorithms, such as federated reinforcement learning (FedRL), federated meta-learning (FedMeta), and federated differential privacy (FedDP), that can learn from diverse and non-independent and identically distributed (non-IID, which refers to the fact that data from various sources at nearby geo-locations are not statistically independent and may follow different probability distributions) data sources while ensuring fairness, diversity, and privacy [139,151]. As more systems are integrated in federated learning, it is becoming crucial to manage the heterogeneity of clients with different computation and communication capabilities, also known as stragglers [167–169].

Another future research direction in FL for air quality monitoring and forecasting is to investigate the role of blockchain and edge computing in improving the scalability and reliability of FL systems. The convergence of blockchain and federated learning in multi-access edge computing can provide a decentralized and tamper-proof ledger that enables secure and transparent data sharing and collaboration among multiple parties in FL. Edge computing, on the other hand, can provide efficient and low-latency computation and storage resources that can reduce the communication and computation overhead of FL. Researchers can explore the integration of blockchain and edge computing with FL to design a secure, efficient, and reliable FL framework for air quality monitoring and forecasting in smart cities [141,170–173].

In addition to FL, another promising solution for air quality monitoring and forecasting is multi-task learning (MTL), which is a machine learning technique that enables multiple tasks to be learned simultaneously by sharing common knowledge across tasks. Future research directions in MTL for air quality monitoring and forecasting can focus on developing more effective and scalable MTL algorithms that can handle large-scale and heterogeneous data, as well as investigating the impact of task relationships and data distribution on the performance of MTL models [104]. Another future research direction in air quality monitoring and forecasting is to integrate physical models with machine learning models to combine the strengths of both approaches and provide more accurate, interpretable, and robust predictions of air quality in smart cities [83].

Finally, a key challenge in air quality monitoring and forecasting is the lack of high-quality and comprehensive data, particularly in developing countries and rural areas. Future research can focus on developing innovative data collection and fusion techniques at high resolution scales and incorporating other sources of pollution-related data, such as satellite remote sensing of environmental pollutants in the atmosphere and monitoring of human activities and social media. These sources can complement the traditional air quality monitoring networks and provide a more holistic and fine-grained understanding of air quality dynamics in smart cities generally [133].

In conclusion, the future research directions and potential solutions for advancing air quality monitoring and forecasting using FL, MTL, and the integration of physical models and machine learning models are diverse and multifaceted. However, addressing the challenges and limitations of these approaches will require interdisciplinary collaborations and innovative methodologies that can bridge the gap between theory and practice. Ultimately, these efforts can lead to more sustainable and healthy smart cities.

6. Conclusions

This paper emphasizes the significance of federated learning (FL) and multi-access edge computing (MEC) approaches in the realm of air quality monitoring and forecasting, particularly within the context of smart environments and cities. Drawing upon a thorough literature review and in-depth analysis of the state-of-the-art techniques, we highlight the growing interest and emerging trends, as well as the potential benefits and constraints of these technologies empowered with deep machine learning paradigms. With the rapid deployment and technological advances of combinations of new generations of wireless mobile networks (5G and beyond) with MEC technology, where computing resources and processing power are pushed closer to data sources, several applications, including air quality monitoring and control, will be offered with more reliable and low-latency connectivity, leading to fast processing and improved real-time analytics and machine learning that make it possible to gain more insights into and value from the generated data. At the same time, federated learning, as a distributed and collaborative learning approach, supports cooperative model training across various edge devices without centralizing data, enhancing data privacy and security. It offers smarter solutions that can handle the massive volume of heterogeneous data from diverse sources and IoT devices at the network edge. Moreover, it balances the edge–cloud interplay to give the models better performance, reliability, privacy preservation, generalizability and interpretability. The capabilities of federated learning, as elucidated in this review, can hold significant implications for pollution control, environmental protection, and public health interventions. They offer a path towards a more comprehensive and fine-grained real-time understanding of air quality dynamics, which can aid in creating more sustainable and healthy smart cities.

Depending on the scale and purpose of monitoring, deployment can range from simple measuring equipment employed in private indoor areas to massive equipment for ambient systems in large urban and industrial areas or even around the globe. Various strategies are proposed in the literature to combine MEC and FL, reducing reliance on centralized systems. Systems like "AirSense", "SmartAir", and HAPADS can be empowered with edge computing and federated learning for real-time air quality measurements. There is also a growing interest in using advanced machine learning methods, including stacking, boosting, ensemble learning, and hybrid models, in air quality monitoring. Ensemble learning, which combines predictions from different models, has shown potential in understanding air quality, especially when using varied data sources. Boosting corrects earlier models' errors, while hybrid models combine deep learning with traditional time-series approaches. It is important to recognize that the success of each method depends on the specific use case and type of data. As research continues, understanding the benefits and challenges of these techniques will be important for improving air quality monitoring systems.

Further research should be directed towards addressing the limitations and leveraging the potential of federated learning and mobile edge computing in building more effective and robust air quality monitoring and intelligent decision-making ecosystems and, in turn, protecting public health in areas such as workplaces, offices, malls, and schools. While this study did not delve deeply into the integration of physical models with machine learning models or the development of innovative data collection and fusion techniques, these may be potential areas for future exploration based on the strengths of federated learning identified herein. In addition, this paper did not incorporate empirical evaluations of the proposed solutions but relied on a review of the existing literature. Addressing these limitations by conducting comparative evaluations of different machine learning paradigms and testing their effectiveness in real-world scenarios is another promising direction for future research.

Author Contributions: Conceptualization, E.-S.M.E.-A. and S.A.; methodology, E.-S.M.E.-A. and S.A.; software, D.S.; validation, S.S., S.A. and E.-S.M.E.-A.; formal analysis, S.S., E.-S.M.E.-A. and S.A.; investigation, Y.-S.C. and S.A.; resources, Y.-S.C., S.S. and S.A.; data curation, Y.-S.C., S.S., E.-S.M.E.-A. and D.S.; writing—original draft preparation, Y.-S.C., D.S., S.H. and S.A.; writing—review and editing, S.A., S.H. and E.-S.M.E.-A.; visualization, D.S.; supervision, S.A. and J.G.B.; project administration, S.A. and J.G.B.; funding acquisition, J.G.B. All authors have read and agreed to the published version of the manuscript.

Funding: This publication emanated from research supported in part by grants from Science Foundation Ireland (SFI) under grant numbers 16/RC/3918 (Confirm) and 12/RC/2289_P2 (Insight) and in part by a grant from SDAIA-KFUPM JRC-AI-REP-04. For the purpose of open access, the authors have applied a CC BY public copyright license to any accepted manuscript arising from this submission.

Institutional Review Board Statement: No human-related data to declare.

Data Availability Statement: The authors declare no data to share.

Acknowledgments: The authors would like to thank their universities and research centers for providing resources during this work. The second author would also like to acknowledge the fellowship at SDAIA-KFUPM Joint Research Center for Artificial Intelligence under grant number JRC-AI-RFP-04.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brook , R.D.; Rajagopalan, S.; Pope, C.A., III; Brook, J.R.; Bhatnagar, A.; Diez-Roux, A.V.; Holguin, F.; Hong, Y.; Luepker, R.V.; Mittleman, M.A.; et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **2010**, *121*, 2331–2378. [CrossRef]
2. Dockery, D.W.; Pope, C.A.; Xu, X.; Spengler, J.D.; Ware, J.H.; Fay, M.E.; Ferris, B.G., Jr.; Speizer, F.E. An association between air pollution and mortality in six US cities. *N. Engl. J. Med.* **1993**, *329*, 1753–1759. [CrossRef]
3. Pope, C.A.; Burnett, R.T.; Thun, M.J.; Calle, E.E.; Krewski, D.; Ito, K.; Thurston, G.D. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J. Am. Med. Assoc.* **2002**, *287*, 1132–1141. [CrossRef]
4. Li, Y.; Guo, J.e.; Sun, S.; Li, J.; Wang, S.; Zhang, C. Air quality forecasting with artificial intelligence techniques: A scientometric and content analysis. *Environ. Model. Softw.* **2022**, *149*, 105329. [CrossRef]
5. Hernandez, D.; Cano, J.C.; Silla, F.; Calafate, C.T.; Cecilia, J.M. AI-Enabled Autonomous Drones for Fast Climate Change Crisis Assessment. *IEEE Internet Things J.* **2022**, *9*, 7286–7297. [CrossRef]
6. Peyman, M.; Copado, P.J.; Tordecilla, R.D.; Martins, L.D.C.; Xhafa, F.; Juan, A.A. Edge computing and iot analytics for agile optimization in intelligent transportation systems. *Energies* **2021**, *14*, 6309. [CrossRef]

7. Hamm, A.; Willner, A.; Schieferdecker, I. Edge computing: A comprehensive survey of current initiatives and a roadmap for a sustainable edge computing development. *arXiv* **2019**, arXiv:1912.08530.
8. Tao, C.; Jia, M.; Wang, G.; Zhang, Y.; Zhang, Q.; Wang, X.; Wang, Q.; Wang, W. Time-sensitive prediction of NO_2 concentration in China using an ensemble machine learning model from multi-source data. *J. Environ. Sci.* **2023**, *137*, 30–40. [CrossRef]
9. Shen, Y.; Wu, T.; Lian, A.; Gao, J.; Peng, F.; Song, G.; Wu, X.; Cui, Y.; Liu, X.; Wan, Y.; et al. Dynamic emission characteristics and control strategies of air pollutants from motor vehicles in downtown Beijing, China. *J. Environ. Sci.* **2024**, *136*, 637–646. [CrossRef]
10. Jabbar, W.A.; Subramaniam, T.; Ong, A.E.; Shu'Ib, M.I.; Wu, W.; de Oliveira, M.A. LoRaWAN-Based IoT System Implementation for Long-Range Outdoor Air Quality Monitoring. *Internet Things* **2022**, *19*, 100540. [CrossRef]
11. Zhu, Y.; Al-Ahmed, S.A.; Shakir, M.Z.; Olszewska, J.I. LSTM-Based IoT-Enabled CO_2 Steady-State Forecasting for Indoor Air Quality Monitoring. *Electronics* **2022**, *12*, 107. [CrossRef]
12. Levy, J.I.; Hammitt, J.K.; Spengler, J.D. Estimating the mortality impacts of particulate matter: What can be learned from between-study variability? *Environ. Health Perspect.* **2000**, *108*, 109–117. [CrossRef] [PubMed]
13. Truong, T.P.; Nguyen, D.T.; Truong, P.V. Design and deployment of an IoT-based air quality monitoring system. *Int. J. Environ. Sci. Dev.* **2021**, *12*, 139–145. [CrossRef]
14. Sung, Y.; Lee, S.; Kim, Y.; Park, H. Development of a Smart Air Quality Monitoring System and its Operation. *Asian J. Atmos. Environ. (Ajae)* **2019**, *13*, 30–38. [CrossRef]
15. Sung, W.T.; Hsiao, S.J. Building an indoor air quality monitoring system based on the architecture of the Internet of Things. *Eurasip J. Wirel. Commun. Netw.* **2021**, *2021*, 153. [CrossRef]
16. Kaur, S.; Sharma, S.; Bawa, S. Smart indoor air quality monitoring system. *Int. J. Recent Technol. Eng.* **2019**, *8*, 989–996.
17. Saini, J.; Dutta, M.; Marques, G. Indoor air quality monitoring systems based on internet of things: A systematic review. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4942. [CrossRef]
18. Marques, G.; Saini, J.; Dutta, M.; Singh, P.K.; Hong, W.C. Indoor air quality monitoring systems for enhanced living environments: A review toward sustainable smart cities. *Sustainability* **2020**, *12*, 4024. [CrossRef]
19. Bishop, S. Air Quality Monitoring 2.0: How Different Types of Air Monitoring Technologies Are Contributing to a More Holistic Understanding of Air Pollution. Available online: https://www.clarity.io/blog/air-quality-monitoring-2-0-how-different-types-of-air-monitoring-technologies-are-contributing-to-a-more-holistic-understanding-of-air-pollution (accessed on 14 April 2023).
20. Sayeed, A.; Choi, Y.; Eslami, E.; Jung, J.; Lops, Y.; Salman, A.K.; Lee, J.B.; Park, H.J.; Choi, M.H. A novel CMAQ-CNN hybrid model to forecast hourly surface-ozone concentrations 14 days in advance. *Sci. Rep.* **2021**, *11*, 10891. [CrossRef]
21. Baklanov, A.; Zhang, Y. Advances in air quality modeling and forecasting. *Glob. Transit.* **2020**, *2*, 261–270. [CrossRef]
22. Pitiranggon, M.; Johnson, S.; Haney, J.; Eisl, H.; Ito, K. Long-term trends in local and transported PM2. 5 pollution in New York City. *Atmos. Environ.* **2021**, *248*, 118238. [CrossRef]
23. Jain, R.; Shah, H. An anomaly detection in smart cities modeled as wireless sensor network. In Proceedings of the IEEE International Conference on Signal and Information Processing (IConSIP), Nanded, India, 6–8 October 2016; pp. 1–5.
24. Trinh, H.D.; Giupponi, L.; Dini, P. Urban Anomaly Detection by processing Mobile Traffic Traces with LSTM Neural Networks. In Proceedings of the Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks Workshops, Boston, MA, USA, 10–13 June 2019. [CrossRef]
25. Rollo, F.; Bachechi, C.; Po, L. Anomaly detection and repairing for improving air quality monitoring. *Sensors* **2023**, *23*, 640. [CrossRef]
26. Kulshrestha, U.; Durbha, S. Edge analytics and complex event processing for real time air pollution monitoring and control. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 893–896.
27. Dutta, A.K.; Sampson, J.; Ahmad, S.; Avudaiappan, T.; Narayanasamy, K.; Pustokhina, I.V.; Pustokhin, D.A. Hybrid Deep Learning Enabled Air Pollution Monitoring in ITS Environment. *Comput. Mater. Contin.* **2022**, *72*, 1157–1172.
28. McCreddin, A.; Alam, M.; McNabola, A. Modelling personal exposure to particulate air pollution: An assessment of time-integrated activity modelling, Monte Carlo simulation & artificial neural network approaches. *Int. J. Hyg. Environ. Health* **2015**, *218*, 107–116. [PubMed]
29. Husaini, D.C.; Reneau, K.; Balam, D. Air pollution and public health in Latin America and the Caribbean (LAC): A systematic review with meta-analysis. *Beni-Suef Univ. J. Basic Appl. Sci.* **2022**, *11*, 122. [CrossRef] [PubMed]
30. Manisalidis, I.; Stavropoulou, E.; Stavropoulos, A.; Bezirtzoglou, E. Environmental and health impacts of air pollution: A review. *Front. Public Health* **2020**, *8*, 14. [CrossRef]
31. Dave, R.; Seliya, N.; Siddiqui, N. The Benefits of Edge Computing in Healthcare, Smart Cities, and IoT. *J. Comput. Sci. Appl.* **2021**, *9*, 23–34. [CrossRef]
32. Abimannan, S.; Ramadoss, R.; Elango, N.; Hsu, C.H. EMAPM: Enterprise Modernization Autonomic Predictive Model in Hybrid Cloud Environments. In Proceedings of the 2017 IEEE 7th International Symposium on Cloud and Service Computing, SC2 2017, Kanazawa, Japan, 22–25 November 2017.
33. Durresi, M.; Subashi, A.; Durresi, A.; Barolli, L.; Uchida, K. Secure communication architecture for internet of things using smartphones and multi-access edge computing in environment monitoring. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 1631–1640. [CrossRef]

34. Nguyen, D.V.; Zettsu, K. Spatially-distributed Federated Learning of Convolutional Recurrent Neural Networks for Air Pollution Prediction. In Proceedings of the IEEE International Conference on Big Data, Big Data, Orlando, FL, USA, 15–18 December 2021.
35. Pham, Q.V.; Ruby, R.; Fang, F.; Nguyen, D.C.; Yang, Z.; Le, M.; Ding, Z.; Hwang, W.J. Aerial Computing: A New Computing Paradigm, Applications, and Challenges. *IEEE Internet Things J.* **2022**, *9*, 8339–8363. [CrossRef]
36. Song, M.; Lee, Y.; Kim, K. Reward-Oriented Task Offloading under Limited Edge Server Power for Multiaccess Edge Computing. *IEEE Internet Things J.* **2021**, *8*, 13425–13438. [CrossRef]
37. Xia, S.; Yao, Z.; Wu, G.; Li, Y. Distributed Offloading for Cooperative Intelligent Transportation Under Heterogeneous Networks. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 16701–16714. [CrossRef]
38. Amiri, M.M.; Gündüz, D. Machine learning at the wireless edge: Distributed stochastic gradient descent over-the-air. *IEEE Trans. Signal Process.* **2020**, *68*, 2155–2169. [CrossRef]
39. Lim, W.Y.B.; Luong, N.C.; Hoang, D.T.; Jiao, Y.; Liang, Y.C.; Yang, Q.; Niyato, D.; Miao, C. Federated learning in mobile edge networks: A comprehensive survey. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 2031–2063. [CrossRef]
40. Sun, W.; Zhao, Y.; Ma, W.; Guo, B.; Xu, L.; Duong, T.Q. Accelerating Convergence of Federated Learning in MEC with Dynamic Community. *IEEE Trans. Mob. Comput.* **2023**. [CrossRef]
41. Gadekallu, T.R.; Pham, Q.V.; Huynh-The, T.; Bhattacharya, S.; Maddikunta, P.K.R.; Liyanage, M. Federated learning for big data: A survey on opportunities, applications, and future directions. *arXiv* **2021**, arXiv:2110.04160.
42. Beltrán, E.T.M.; Pérez, M.Q.; Sánchez, P.M.S.; Bernal, S.L.; Bovet, G.; Pérez, M.G.; Pérez, G.M.; Celdrán, A.H. Decentralized federated learning: Fundamentals, state-of-the-art, frameworks, trends, and challenges. *arXiv* **2022**, arXiv:2211.08413.
43. Hu, B.; Gao, Y.; Liu, L.; Ma, H. Federated Region-Learning: An Edge Computing Based Framework for Urban Environment Sensing. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 9–13 December 2018.
44. Reinnes, J. Investigating and Developing Efficient Federated Learning for Air Pollution Monitoring. Master's Thesis, UiT Norges Arktiske Universitet, Tromsø, Norway, 2022 .
45. Gao, Y.; Liu, L.; Hu, B.; Lei, T.; Ma, H. Federated Region-Learning for Environment Sensing in Edge Computing System. *IEEE Trans. Netw. Sci. Eng.* **2020**, *7*, 2192–2204. [CrossRef]
46. Neo, E.X.; Hasikin, K.; Mokhtar, M.I.; Lai, K.W.; Azizan, M.M.; Razak, S.A.; Hizaddin, H.F. Towards Integrated Air Pollution Monitoring and Health Impact Assessment Using Federated Learning: A Systematic Review. *Front. Public Health* **2022**, *10*, 851553. [CrossRef]
47. Chhikara, P.; Tekchandani, R.; Kumar, N.; Guizani, M.; Hassan, M.M. Federated learning and autonomous UAVs for hazardous zone detection and AQI prediction in IoT environment. *IEEE Internet Things J.* **2021**, *8*, 15456–15467. [CrossRef]
48. Gao, Y.; Liu, L.; Zheng, X.; Zhang, C.; Ma, H. Federated sensing: Edge-cloud elastic collaborative learning for intelligent sensing. *IEEE Internet Things J.* **2021**, *8*, 11100–11111. [CrossRef]
49. Vecchio, M.; Azzoni, P.; Menychtas, A.; Maglogiannis, I.; Felfernig, A. A fully open-source approach to intelligent edge computing: Agile's lesson. *Sensors* **2021**, *21*, 1309. [CrossRef]
50. Ali, B.; Gregory, M.A.; Li, S. Multi-access edge computing architecture, data security and privacy: A review. *IEEE Access* **2021**, *9*, 18706–18721. [CrossRef]
51. Joshi, K.; Anandaram, H.; Khanduja, M.; Kumar, R.; Saini, V.; Mohialden, Y.M. Recent Challenges on Edge AI with Its Application: A Brief Introduction. In *Studies in Computational Intelligence*; Springer: Cham, Switzerland, 2023; Volume 1072.
52. Braeken, A.; Liyanage, M. Highly efficient key agreement for remote patient monitoring in MEC-enabled 5G networks. *J. Supercomput.* **2021**, *77*, 5562–5583. [CrossRef]
53. Kristiani, E.; Yang, C.T.; Huang, C.Y.; Wang, Y.T.; Ko, P.C. The Implementation of a Cloud-Edge Computing Architecture Using OpenStack and Kubernetes for Air Quality Monitoring Application. *Mob. Netw. Appl.* **2021**, *26*, 1070–1092. [CrossRef]
54. Su, X.; Liu, X.; Motlagh, N.H.; Cao, J.; Su, P.; Pellikka, P.; Liu, Y.; Petaja, T.; Kulmala, M.; Hui, P.; et al. Intelligent and Scalable Air Quality Monitoring with 5G Edge. *IEEE Internet Comput.* **2021**, *25*, 35–44. [CrossRef]
55. Al-Quraan, M.; Mohjazi, L.; Bariah, L.; Centeno, A.; Zoha, A.; Arshad, K.; Assaleh, K.; Muhaidat, S.; Debbah, M.; Imran, M.A. Edge-native intelligence for 6G communications driven by federated learning: A survey of trends and challenges. *IEEE Trans. Emerg. Top. Comput. Intell.* **2023**, 957–959. [CrossRef]
56. Yuan, X.; Chen, J.; Yang, J.; Zhang, N.; Yang, T.; Han, T.; Taherkordi, A. FedSTN: Graph Representation Driven Federated Learning for Edge Computing Enabled Urban Traffic Flow Prediction. *IEEE Trans. Intell. Transp. Syst.* **2022**, 8738–8748. [CrossRef]
57. Qolomany, B.; Ahmad, K.; Al-Fuqaha, A.; Qadir, J. Particle swarm optimized federated learning for industrial IoT and smart city services. In Proceedings of the IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; pp. 1–6.
58. Polymeni, S.; Skoutas, D.N.; Kormentzas, G.; Skianis, C. FINDEAS: A FinTech-Based Approach on Designing and Assessing IoT Systems. *IEEE Internet Things J.* **2022**, *9*, 25196–25206. [CrossRef]
59. Dey, S.; Pal, S. Federated learning-based air quality prediction for smart cities using BGRU model. In Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM, Sydney, Australia, 14 October 2022. [CrossRef]
60. Van Poppel, M.; Schneider, P.; Peters, J.; Yatkin, S.; Gerboles, M.; Matheeussen, C.; Bartonova, A.; Davila, S.; Signorini, M.; Vogt, M.; et al. SensEURCity: A multi-city air quality dataset collected for 2020/2021 using open low-cost sensor systems. *Sci. Data* **2023**, *10*, 322. [CrossRef]

61. Hua, H.; Li, Y.; Wang, T.; Dong, N.; Li, W.; Cao, J. Edge Computing with Artificial Intelligence: A Machine Learning Perspective. *ACM Comput. Surv.* **2023**, *55*, 1–35. [CrossRef]
62. Putra, K.T.; Chen, H.C.; Prayitno.; Ogiela, M.R.; Chou, C.L.; Weng, C.E.; Shae, Z.Y. Federated compressed learning edge computing framework with ensuring data privacy for pm2.5 prediction in smart city sensing applications. *Sensors* **2021**, *21*, 4586. [CrossRef] [PubMed]
63. Moursi, A.S.; El-Fishawy, N.; Djahel, S.; Shouman, M.A. An IoT enabled system for enhanced air quality monitoring and prediction on the edge. *Complex Intell. Syst.* **2021**, *7*, 2923–2947. [CrossRef] [PubMed]
64. Biondi, K.; Al-Masri, E.; Baiocchi, O.; Jeyaraman, S.; Pospisil, E.; Boyer, G.; De Souza, C.P. Air Pollution Detection System using Edge Computing. In Proceedings of the IEEE International Conference on Engineering Applications, ICEA, Sao Miguel, Portugal, 8–11 July 2019.
65. Xu, Z.; Ran, Y.; Rao, Z. Design and integration of air pollutants monitoring system for emergency management in construction site based on BIM and edge computing. *Build. Environ.* **2022**, *211*, 108725. [CrossRef]
66. Kim, M.; Kim, T.; Park, S.; Lee, K. An Indoor Multi-Environment Sensor System Based on Intelligent Edge Computing. *Electronics* **2023**, *12*, 137. [CrossRef]
67. Idrees, Z.; Zou, Z.; Zheng, L. Edge computing based IoT architecture for low cost air pollution monitoring systems: A comprehensive system analysis, design considerations & development. *Sensors* **2018**, *18*, 3021.
68. Wang, B.; Li, M.; Jin, X.; Guo, C. A Reliable IoT Edge Computing Trust Management Mechanism for Smart Cities. *IEEE Access* **2020**, *8*, 46373–46399. [CrossRef]
69. Tanaka, H.; Yoshida, M.; Mori, K.; Takahashi, N. Multi-access edge computing: A survey. *J. Inf. Process.* **2018**, *26*, 87–97. [CrossRef]
70. Gandotra, P.; Lall, B. Evolving Air Pollution Monitoring Systems for Green 5G: From Cloud to Edge. In Proceedings of the IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), Noida, India, 4–5 June 2020.
71. Marjanovic, M.; Antonic, A.; Zarko, I.P. Edge computing architecture for mobile crowdsensing. *IEEE Access* **2018**, *6*, 10662–10674. [CrossRef]
72. Zhang, Q.; Luo, Y.; Jiang, H.; Zhang, K. Aerial Edge Computing: A Survey. *IEEE Internet Things J.* **2023**. [CrossRef]
73. Peltonen, E.; Ahmad, I.; Aral, A.; Capobianco, M.; Ding, A.Y.; Gil-Castiñeira, F.; Gilman, E.; Harjula, E.; Jurmu, M.; Karvonen, T.; et al. The Many Faces of Edge Intelligence. *IEEE Access* **2022**, *10*, 104769–104782. [CrossRef]
74. Peltonen, E.; Bennis, M.; Capobianco, M.; Debbah, M.; Ding, A.; Gil-Castiñeira, F.; Jurmu, M.; Karvonen, T.; Kelanti, M.; Kliks, A.; et al. 6G white paper on edge intelligence. *arXiv* **2020**, arXiv:2004.14850.
75. Yang, B.; Cao, X.; Bassey, J.; Li, X.; Qian, L. Computation Offloading in Multi-Access Edge Computing: A Multi-Task Learning Approach. *IEEE Trans. Mob. Comput.* **2021**, *20*, 2745–2762. [CrossRef]
76. Khan, I.; Guerrieri, A.; Spezzano, G.; Vinci, A. Occupancy prediction in buildings: An approach leveraging LSTM and Federated Learning. In Proceedings of the IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), Falerna, Italy, 12–15 September 2022; pp. 1–7.
77. Böhm, S.; Wirtz, G. Cloud-Edge Orchestration for Smart Cities: A Review of Kubernetes-based Orchestration Architectures. *EAI Endorsed Trans. Smart Cities* **2022**, *6*, e2. [CrossRef]
78. Wang, S.; Huang, Y.; Clerckx, B. Dynamic Air-Ground Collaboration for Multi-Access Edge Computing. In Proceedings of the IEEE International Conference on Communications, Seoul, Republic of Korea, 16–20 May 2022; pp. 5365–5371. [CrossRef]
79. Popović, I.; Radovanovic, I.; Vajs, I.; Drajic, D.; Gligorić, N. Building Low-Cost Sensing Infrastructure for Air Quality Monitoring in Urban Areas Based on Fog Computing. *Sensors* **2022**, *22*, 1026. [CrossRef] [PubMed]
80. Weiss, M.B.; Gavras, A.; Salva-Garcia, P.; Alcaraz-Calero, J.M.; Wang, Q. Network Management-Edge and Cloud Computing the SliceNet Case. In Proceedings of the 2020 IEEE 17th Annual Consumer Communications and Networking Conference, CCNC, Las Vegas, NV, USA, 10–13 January 2020. [CrossRef]
81. Yan, R.; Liao, J.; Yang, J.; Sun, W.; Nong, M.; Li, F. Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. *Expert Syst. Appl.* **2021**, *169*, 114513. [CrossRef]
82. Pandya, S.; Srivastava, G.; Jhaveri, R.; Babu, M.R.; Bhattacharya, S.; Maddikunta, P.K.R.; Mastorakis, S.; Piran, M.J.; Gadekallu, T.R. Federated learning for smart cities: A comprehensive survey. *Sustain. Energy Technol. Assess.* **2023**, *55*, 102987. [CrossRef]
83. Zhou, X.; Liu, X.; Lan, G.; Wu, J. Federated conditional generative adversarial nets imputation method for air quality missing data. *Knowl.-Based Syst.* **2021**, *228*, 107261. [CrossRef]
84. Agarwal, P.; Sharma, S.; Matta, P. Federated Learning in Intelligent Traffic Management System. In Proceedings of the 2023 Winter Summit on Smart Computing and Networks (WiSSCoN), Chennai, India, 15–17 March 2023; pp. 1–6.
85. Le, D.D.; Tran, A.K.; Dao, M.S.; Nguyen-Ly, K.C.; Le, H.S.; Nguyen-Thi, X.D.; Pham, T.Q.; Nguyen, V.L.; Nguyen-Thi, B.Y. Insights into multi-model federated learning: An advanced approach for air quality index forecasting. *Algorithms* **2022**, *15*, 434. [CrossRef]
86. Liu, Y.; Nie, J.; Li, X.; Ahmed, S.H.; Lim, W.Y.B.; Miao, C. Federated Learning in the Sky: Aerial-Ground Air Quality Sensing Framework with UAV Swarms. *IEEE Internet Things J.* **2021**, *8*, 9827–9837. [CrossRef]
87. Xiang, K.; Zeng, Y. Edge-Driven On-Line IoT Data Search Based on Knowledge Graph Convolutional Networks. *Ssrn Electron. J.* **2023**.

88. Chhikara, P.; Tekchandani, R.; Kumar, N.; Tanwar, S.; Rodrigues, J.J. Federated learning for air quality index prediction using UAV swarm networks. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 7–11 December 2021; pp. 1–6.
89. Chinchole, U.P.; Raut, S. Federated learning for estimating air quality. In Proceedings of the IEEE 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 6–8 July 2021; pp. 1–7.
90. Muruganandam, N.S.; Arumugam, U. Seminal Stacked Long Short-Term Memory (SS-LSTM) Model for Forecasting Particulate Matter (PM2.5 and PM10). *Atmosphere* **2022**, *13*, 1726. [CrossRef]
91. Le, D.D.; Tran, A.K.; Dao, M.S.; Nazmudeen, M.S.H.; Mai, V.T.; Su, N.H. Federated Learning for Air Quality Index Prediction: An Overview. In Proceedings of the International Conference on Knowledge and Systems Engineering, KSE, Nha Trang, Vietnam, 19–21 October 2022; Volume 2022.
92. Huang, G.; Zhao, X.; Lu, Q. A new cross-domain prediction model of air pollutant concentration based on secure federated learning and optimized LSTM neural network. *Environ. Sci. Pollut. Res.* **2023**, *30*, 5103–5125. [CrossRef]
93. Wang, A.; Zha, Z.; Guo, Y.; Chen, S. Software-Defined Networking Enhanced Edge Computing: A Network-Centric Survey. *Proc. IEEE* **2019**, *107*, 1500–1519. [CrossRef]
94. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19. [CrossRef]
95. Huang, Z.; Balasubramanian, B.; Alsudais, A.; Joshi, K. An edge-facilitated message broker for scalable device discovery. In Proceedings of the Second ACM/IEEE Symposium on Edge Computing, San Jose, CA, USA, 12–14 October 2017. [CrossRef]
96. Poongodi, M.; Bourouis, S.; Ahmed, A.N.; Vijayaragavan, M.; Venkatesan, K.G.; Alhakami, W.; Hamdi, M. A Novel Secured Multi-Access Edge Computing based VANET with Neuro fuzzy systems based Blockchain Framework. *Comput. Commun.* **2022**, *192*, 45–86. [CrossRef]
97. Khalfoun, B.; Ben Mokhtar, S.; Bouchenak, S.; Nitu, V. EDEN: Enforcing Location Privacy through Re-identification Risk Assessment: A Federated Learning Approach. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2021**, *5*, 1–25. [CrossRef]
98. Kamble, A.; Champrasert, P. Using Photo Images with Deep Residual Network for PM2.5 Value Estimation. In Proceedings of the Lecture Notes in Networks and Systems, Online, 20 October 2021; Volume 343.
99. Jiang, J.C.; Kantarci, B.; Oktug, S.; Soyata, T. Federated learning in smart city sensing: Challenges and opportunities. *Sensors* **2020**, *20*, 6230. [CrossRef] [PubMed]
100. Gupta, S.; Naidu, P.C.B.; Kuppan, V.; Alagumeenaakshi, M.; Niruban, R.; Swaminathan, J.N. Analysis of Various Toxic Gas Levels Using 5G ML-IoT for Air Quality Monitoring and Forecasting. In *IOT with Smart Systems, Proceedings of the ICTIS 2022, Ahmedabad, India, 22–23 April 2022* ; Smart Innovation, Systems and Technologies; Springer Science and Business Media Deutschland GmbH: Berlin/Heidelberg, Germany, 2023; Volume 312, pp. 789–802. [CrossRef]
101. Hamza, M.A.; Shaiba, H.; Marzouk, R.; Alhindi, A.; Asiri, M.M.; Yaseen, I.; Motwakel, A.; Rizwanullah, M. Big Data Analytics with Artificial Intelligence Enabled Environmental Air Pollution Monitoring Framework. *Comput. Mater. Contin.* **2022**, *73*, 3235–3250.
102. Zhang, C.; Zhou, G.; Li, J.; Chang, F.; Ding, K.; Ma, D. A multi-access edge computing enabled framework for the construction of a knowledge-sharing intelligent machine tool swarm in Industry 4.0. *J. Manuf. Syst.* **2023**, *66*, 56–70. [CrossRef]
103. Wu, Z.; Zhao, W.; Lv, Y. An ensemble LSTM-based AQI forecasting model with decomposition-reconstruction technique via CEEMDAN and fuzzy entropy. *Air Qual. Atmos. Health* **2022**, *15*. [CrossRef]
104. Lyu, Y.; Ju, Q.; Lv, F.; Feng, J.; Pang, X.; Li, X. Spatiotemporal variations of air pollutants and ozone prediction using machine learning algorithms in the Beijing-Tianjin-Hebei region from 2014 to 2021. *Environ. Pollut.* **2022**, *306*, 119420. [CrossRef] [PubMed]
105. Liu, Y.; Zhou, Y.; Lu, J. Exploring the relationship between air pollution and meteorological conditions in China under environmental governance. *Sci. Rep.* **2020**, *10*, 14518. [CrossRef] [PubMed]
106. Zaytar, M.A.; El Amrani, C. Machine Learning Methods for Air Quality Monitoring. In Proceedings of the 3rd International Conference on Networking, Information Systems & Security, Marrakech, Morocco, 31 March–2 April 2020; ACM International Conference Proceeding Series; Association for Computing Machinery: New York, NY, USA, 2020. [CrossRef]
107. Zalakeviciute, R.; Bastidas, M.; Buenaño, A.; Rybarczyk, Y. A Traffic-based method to predict and map urban air quality. *Appl. Sci.* **2020**, *10*, 2035. [CrossRef]
108. Mihăiţă, A.S.; Dupont, L.; Chery, O.; Camargo, M.; Cai, C. Evaluating air quality by combining stationary, smart mobile pollution monitoring and data-driven modelling. *J. Clean. Prod.* **2019**, *221*, 398–418. [CrossRef]
109. Alvear-Puertas, V.E.; Burbano-Prado, Y.A.; Rosero-Montalvo, P.D.; Tözün, P.; Marcillo, F.; Hernandez, W. Smart and Portable Air-Quality Monitoring IoT Low-Cost Devices in Ibarra City, Ecuador. *Sensors* **2022**, *22*, 7015. [CrossRef] [PubMed]
110. Tancev, G. Relevance of drift components and unit-to-unit variability in the predictive maintenance of low-cost electrochemical sensor systems in air quality monitoring. *Sensors* **2021**, *21*, 3298. [CrossRef]
111. Goh, C.C.; Kamarudin, L.M.; Zakaria, A.; Nishizaki, H.; Ramli, N.; Mao, X.; Zakaria, S.M.M.S.; Kanagaraj, E.; Sukor, A.S.A.; Elham, M.F. Real-time in-vehicle air quality monitoring system using machine learning prediction algorithm. *Sensors* **2021**, *21*, 4956. [CrossRef]
112. Chojer, H.; Branco, P.T.; Martins, F.G.; Alvim-Ferraz, M.C.; Sousa, S.I. Can data reliability of low-cost sensor devices for indoor air particulate matter monitoring be improved?—An approach using machine learning. *Atmos. Environ.* **2022**, *286*, 119251. [CrossRef]
113. Adong, P.; Bainomugisha, E.; Okure, D.; Sserunjogi, R. Applying machine learning for large scale field calibration of low-cost PM 2.5 and PM 10 air pollution sensors. *Appl. Lett.* **2022**, *3*, e76. [CrossRef]

114. Kapoor, N.R.; Kumar, A.; Kumar, A.; Kumar, A.; Mohammed, M.A.; Kumar, K.; Kadry, S.; Lim, S. Machine Learning-Based CO_2 Prediction for Office Room: A Pilot Study. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 9404807. [CrossRef]
115. Lim, C.C.; Kim, H.; Vilcassim, M.J.; Thurston, G.D.; Gordon, T.; Chen, L.C.; Lee, K.; Heimbinder, M.; Kim, S.Y. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea. *Environ. Int.* **2019**, *131*, 105022. [CrossRef] [PubMed]
116. Aram, S.; Nketiah, E.; Saalidong, B.; Wang, H.; Afitiri, A.R.; Akoto, A.; Lartey, P. Machine learning-based prediction of air quality index and air quality grade: A comparative analysis. *Int. J. Environ. Sci. Technol.* **2023**, 1–16. [CrossRef]
117. Guo, S.; Tao, X.; Liang, L. Exploring Natural and Anthropogenic Drivers of PM2.5 Concentrations Based on Random Forest Model: Beijing–Tianjin–Hebei Urban Agglomeration, China. *Atmosphere* **2023**, *14*, 381. [CrossRef]
118. Bamrah, S.K.; Srivatsan, S.; Gayathri, K.S. Region Classification for Air Quality Estimation Using Deep Learning and Machine Learning Approach. In Proceedings of the Lecture Notes in Electrical Engineering, Online, 1 January 2023; Volume 946.
119. Zhu, M.; Xie, J. Investigation of nearby monitoring station for hourly PM2.5 forecasting using parallel multi-input 1D-CNN-biLSTM. *Expert Syst. Appl.* **2023**, *211*, 118707. [CrossRef]
120. Gao, X.; Li, W. A graph-based LSTM model for PM2.5 forecasting. *Atmos. Pollut. Res.* **2021**, *12*, 101150. [CrossRef]
121. Abimannan, S.; Chang, Y.S.; Lin, C.Y. Air Pollution Forecasting Using LSTM-Multivariate Regression Model. In *Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2020; Volume 11894 LNCS, pp. 318–326.
122. Lin, C.Y.; Chang, Y.S.; Abimannan, S. Ensemble multifeatured deep learning models for air quality forecasting. *Atmos. Pollut. Res.* **2021**, *12*, 101045. [CrossRef]
123. Chang, Y.S.; Abimannan, S.; Chiao, H.T.; Lin, C.Y.; Huang, Y.P. An ensemble learning based hybrid model and framework for air pollution forecasting. *Environ. Sci. Pollut. Res.* **2020**, *27*, 38155–38168. [CrossRef]
124. Suriya; Natsagdorj, N.; Aorigele; Zhou, H.; Sachurila. Prediction of PM2.5 concentration in Ulaanbaatar with deep learning models. *Urban Clim.* **2023**, *47*, 101357.
125. Vicente, G.; Marques, G. Air quality monitoring through LoRa technologies: A literature review. In Proceedings of the IEEE International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, 8–9 November 2020; pp. 350–354.
126. Singh, S.; Rathore, S.; Alfarraj, O.; Tolba, A.; Yoon, B. A framework for privacy-preservation of IoT healthcare data using Federated Learning and blockchain technology. *Future Gener. Comput. Syst.* **2022**, *129*, 380–388. [CrossRef]
127. Morris, A.S.; Langari, R. Sensor Technologies. In *Measurements and Instrumentation: Theory and Application*; Academic Press: Oxford, UK, 2016; Chapter 13, pp. 375–405.
128. Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE Signal Process. Mag.* **2020**, *37*, 50–60. [CrossRef]
129. Zhao, Y.; Hu, N.; Zhao, Y.; Zhu, Z. A secure and flexible edge computing scheme for AI-driven industrial IoT. *Clust. Comput.* **2023**, *26*, 283–301. [CrossRef]
130. Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L. Edge computing: Vision and challenges. *IEEE Internet Things J.* **2016**, *3*, 637–646. [CrossRef]
131. Sayeed, A.; Lops, Y.; Choi, Y.; Jung, J.; Salman, A.K. Bias correcting and extending the PM forecast by CMAQ up to 7 days using deep convolutional neural networks. *Atmos. Environ.* **2021**, *253*, 118376. [CrossRef]
132. Ojagh, S.; Cauteruccio, F.; Terracina, G.; Liang, S.H. Enhanced air quality prediction by edge-based spatiotemporal data preprocessing. *Comput. Electr. Eng.* **2021**, *96*, 107572. [CrossRef]
133. Wang, J.; Zhang, X.; Guo, Z.; Lu, H. Developing an early-warning system for air quality prediction and assessment of cities in China. *Expert Syst. Appl.* **2017**, *84*, 102–116. [CrossRef]
134. Yan, C.; Nie, W.; Äijälä, M.; Rissanen, M.P.; Canagaratna, M.R.; Massoli, P.; Junninen, H.; Jokinen, T.; Sarnela, N.; Häme, S.A.; et al. Source characterization of highly oxidized multifunctional compounds in a boreal forest environment using positive matrix factorization. *Atmos. Chem. Phys.* **2016**, *16*, 12715–12731. [CrossRef]
135. Sokhi, R.S.; Moussiopoulos, N.; Baklanov, A.; Bartzis, J.; Coll, I.; Finardi, S.; Friedrich, R.; Geels, C.; Grönholm, T.; Halenka, T.; et al. Advances in air quality research–current and emerging challenges. *Atmos. Chem. Phys.* **2022**, *22*, 4615–4703. [CrossRef]
136. Zhang, Y.; Liu, J.H.; Wang, C.Y.; Wei, H.Y. Decomposable intelligence on cloud-edge iot framework for live video analytics. *IEEE Internet Things J.* **2020**, *7*, 8860–8873. [CrossRef]
137. Wu, Y. Cloud-edge orchestration for the Internet of Things: Architecture and AI-powered data processing. *IEEE Internet Things J.* **2020**, *8*, 12792–12805. [CrossRef]
138. Sun, C.; Wu, X.; Li, X.; Fan, Q.; Wen, J.; Leung, V.C. Cooperative computation offloading for multi-access edge computing in 6G mobile networks via soft actor critic. *IEEE Trans. Netw. Sci. Eng.* **2021**, 1. [CrossRef]
139. Wang, X.; Wang, C.; Li, X.; Leung, V.C.; Taleb, T. Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching. *IEEE Internet Things J.* **2020**, *7*, 9441–9455. [CrossRef]
140. Wen, J.; Zhang, Z.; Lan, Y.; Cui, Z.; Cai, J.; Zhang, W. A survey on federated learning: Challenges and applications. *Int. J. Mach. Learn. Cybern.* **2022**, *14*, 513–535. [CrossRef] [PubMed]
141. Zheng, Z.; Zhou, Y.; Sun, Y.; Wang, Z.; Liu, B.; Li, K. Applications of federated learning in smart cities: Recent advances, taxonomy, and open challenges. *Connect. Sci.* **2022**, *34*, 1–28. [CrossRef]

142. Xianjia, Y.; Queralta, J.P.; Heikkonen, J.; Westerlund, T. Federated learning in robotic and autonomous systems. *Procedia Comput. Sci.* **2021**, *191*, 135–142. [CrossRef]
143. Li, Y.; Tao, X.; Zhang, X.; Liu, J.; Xu, J. Privacy-preserved federated learning for autonomous driving. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 8423–8434. [CrossRef]
144. Maag, B.; Zhou, Z.; Thiele, L. A survey on sensor calibration in air pollution monitoring deployments. *IEEE Internet Things J.* **2018**, *5*, 4857–4870. [CrossRef]
145. Concas, F.; Mineraud, J.; Lagerspetz, E.; Varjonen, S.; Liu, X.; Puolamäki, K.; Nurmi, P.; Tarkoma, S. Low-cost outdoor air quality monitoring and sensor calibration: A survey and critical analysis. *Acm Trans. Sens. Netw. (Tosn)* **2021**, *17*, 1–44. [CrossRef]
146. Chen, L.; Long, H.; Xu, J.; Wu, B.; Zhou, H.; Tang, X.; Peng, L. Deep Citywide Multisource Data Fusion-Based Air Quality Estimation. *IEEE Trans. Cybern.* **2023**, 1–12. [CrossRef]
147. Liang, L.; Daniels, J.; Bailey, C.; Hu, L.; Phillips, R.; South, J. Integrating low-cost sensor monitoring, satellite mapping, and geospatial artificial intelligence for intra-urban air pollution predictions. *Environ. Pollut.* **2023**, *331*, 121832. [CrossRef]
148. Mohindru, G.; Mondal, K.; Banka, H. Internet of Things and data analytics: A current review. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2020**, *10*, 1341. [CrossRef]
149. Wang, W.; Feng, C.; Zhang, B.; Gao, H. Environmental monitoring based on fog computing paradigm and internet of things. *IEEE Access* **2019**, *7*, 127154–127165. [CrossRef]
150. Wang, Y.; Su, Z.; Zhang, N.; Benslimane, A. Learning in the air: Secure federated learning for UAV-assisted crowdsensing. *IEEE Trans. Netw. Sci. Eng.* **2020**, *8*, 1055–1069. [CrossRef]
151. Perifanis, V.; Efraimidis, P.S. Federated neural collaborative filtering. *Knowl.-Based Syst.* **2022**, *242*, 108441. [CrossRef]
152. Soni, M.; Nayak, N.R.; Kalra, A.; Degadwala, S.; Singh, N.K.; Singh, S. Energy efficient multi-tasking for edge computing using federated learning. *Int. J. Pervasive Comput. Commun.* 2022, *ahead-of-print*. [CrossRef]
153. Dogo, E.M.; Salami, A.F.; Aigbavboa, C.O.; Nkonyana, T. Taking cloud computing to the extreme edge: A review of mist computing for smart cities and industry 4.0 in africa. In *EAI/Springer Innovations in Communication and Computing*; Springer: Cham, Switzerland, 2019. [CrossRef]
154. Nimkar, S.; Khanapurkar, M.M. Edge Computing for IoT: A Use Case in Smart City Governance. In Proceedings of the 2021 International Conference on Computational Intelligence and Computing Applications, ICCICA 2021, Nagpur, India, 26–27 November 2021. [CrossRef]
155. Maltezos, E.; Karagiannidis, L.; Dadoukis, A.; Petousakis, K.; Misichroni, F.; Ouzounoglou, E.; Gounaridis, L.; Gounaridis, D.; Kouloumentas, C.; Amditis, A. Public safety in smart cities under the edge computing concept. In Proceedings of the 2021 IEEE International Mediterranean Conference on Communications and Networking, MeditCom 2021, Athens, Greece, 7–10 September 2021; pp. 88–93 . [CrossRef]
156. Alsamhi, S.H.; Shvetsov, A.V.; Kumar, S.; Hassan, J.; Alhartomi, M.A.; Shvetsova, S.V.; Sahal, R.; Hawbani, A. Computing in the Sky: A Survey on Intelligent Ubiquitous Computing for UAV-Assisted 6G Networks and Industry 4.0/5.0. *Drones* **2022**, *6*, 177. [CrossRef]
157. Xia, Z.; Du, J.; Ren, Y.; Han, Z. Distributed Artificial Intelligence Enabled Aerial-Ground Networks: Architecture, Technologies and Challenges. *IEEE Access* **2022**, *10*, 105447–105457. [CrossRef]
158. Putrada, A.G.; Abdurohman, M.; Perdana, D.; Nuha, H.H. EdgeSL: Edge-Computing Architecture on Smart Lighting Control with Distilled KNN for Optimum Processing Time. *IEEE Access* **2023**, *11*, 64697–64712. [CrossRef]
159. Palviainen, M.; Suksi, J. Data marketplace research: A review of the state-of-the-art with a focus on smart cities and on edge data exchange and trade. In Proceedings of the 2023 Smart City Symposium Prague (SCSP), Prague, Czech Republic, 25–26 May 2023; pp. 1–7.
160. Morabito, R.; Cozzolino, V.; Ding, A.Y.; Beijar, N.; Ott, J. Consolidate IoT Edge Computing with Lightweight Virtualization. *IEEE Netw.* **2018**, *32*, 102–111. [CrossRef]
161. Theodorou, V.; Xezonaki, M.E. Network slicing for multi-tenant edge processing over shared IoT infrastructure. In Proceedings of the IEEE Conference on Network Softwarization: Bridging the Gap Between AI and Network Softwarization, NetSoft 2020, Ghent, Belgium, 29 June–3 July 2020. [CrossRef]
162. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [CrossRef]
163. Bejnordi, B.E.; Veta, M.; Van Diest, P.J.; Van Ginneken, B.; Karssemeijer, N.; Litjens, G.; Van Der Laak, J.A.; Hermsen, M.; Manson, Q.F.; Balkenhol, M.; et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *J. Am. Med. Assoc.* **2017**, *318*, 2199–2210. [CrossRef]
164. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the Artificial Intelligence and Statistics, PMLR, Fort Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282.
165. Motlagh, N.H.; Petäjä, T.; Kulmala, M.; Trachoma, S.; Lagerspetz, E.; Nurmi, P.; Li, X.; Varjonen, S.; Mineraud, J.; Siekkinen, M.; et al. Toward Massive Scale Air Quality Monitoring. *IEEE Commun. Mag.* **2020**, *58*, 54–59. [CrossRef]
166. Shi, T.; Cai, Z.; Li, Y. Query Recombination: To Process a Large Number of Concurrent Top-k Queries towards IoT Data on an Edge Server. In Proceedings of the Proceedings—International Conference on Distributed Computing Systems, Bologna, Italy, 10–13 July 2022. [CrossRef]

167. Reisizadeh, A.; Tziotis, I.; Hassani, H.; Mokhtari, A.; Pedarsani, R. Straggler-resilient federated learning: Leveraging the interplay between statistical accuracy and system heterogeneity. *IEEE J. Sel. Areas Inf. Theory* **2022**, *3*, 197–205. [CrossRef]
168. Li, S.; Yao, D.; Liu, J. FedVS: Straggler-Resilient and Privacy-Preserving Vertical Federated Learning for Split Models. *arXiv* **2023**, arXiv:2304.13407.
169. Schlegel, R.; Kumar, S.; Rosnes, E.; i Amat, A.G. CodedPaddedFL and CodedSecAgg: Straggler mitigation and secure aggregation in federated learning. *IEEE Trans. Commun.* **2023**. [CrossRef]
170. Islam, A.; Al Amin, A.; Shin, S.Y. FBI: A federated learning-based blockchain-embedded data accumulation scheme using drones for Internet of Things. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 972–976. [CrossRef]
171. Ur Rehman, M.H.; Salah, K.; Damiani, E.; Svetinovic, D. Towards blockchain-based reputation-aware federated learning. In Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 6–9 July 2020; pp. 183–188.
172. Núñez-Gómez, C.; Caminero, B.; Carrión, C. HIDRA: A distributed blockchain-based architecture for fog/edge computing environments. *IEEE Access* **2021**, *9*, 75231–75251. [CrossRef]
173. Billah, M.; Mehedi, S.T.; Anwar, A.; Rahman, Z.; Islam, R. A systematic literature review on blockchain enabled federated learning framework for internet of vehicles. *arXiv* **2022**, arXiv:2203.05192.