An Investigation of Federated Learning for Air Quality Forecasting and Monitoring

Yuhao Wu

Department of Statistic and Data Science, Capital University of Economics and business, Beijing, China

Yuhaowu0109zz@outlook.com

Abstract. Air quality, as a global issue, impacts people's health and daily life. It requires precise prediction and monitoring for sustainable urban management. Conventional centralized air quality prediction methods are limited by data privacy, high communication costs, and low scalability in distributed environments. Federated learning offers a solution that collaboratively trains global models without sharing raw data. The passage conducted a comprehensive literature analysis to categorize FL applications in air quality research into four technical streams: Probabilistic Graphical Models (Federated Bayesian Networks) for causal inference; Time-series Deep Learning (integrating LSTM/CNN with FedAvg/FedProx) for temporal pattern extraction; Spatio-temporal Graph Neural Networks (GC-LSTM) for capturing complex spatial dependencies; and Multi-model Ensemble with Transfer Learning for heterogeneous client adaptation. Nevertheless, there are significant gaps that remain in terms of model interpretability and the ability to generalize across climate variations. All in all, this article provides a relatively comprehensive overview of the application of federated learning in air quality forecasting and monitoring. It will assist researchers in advancing this field in the future.

Keywords: Federated learning, Air quality forecasting, Air quality monitoring, Spatio-temporal graph neural networks.

1. Introduction

Air pollution has become a severe global environmental and public health problem, contributing to increased morbidity and mortality worldwide. Accurate air quality monitoring and forecasting are essential for environmental governance and the sustainable development of cities. Traditional air quality prediction methods typically rely on centralized machine learning models, in which data collected from distributed monitoring stations are aggregated and processed on a central server. Although such approaches can achieve high predictive performance, they raise critical concerns regarding data privacy, high communication costs, and limited scalability in distributed environments [1].

Federated learning is a suitable solution to address challenges such as data silos, data heterogeneity, and the high cost of data transmission and storage. The standard algorithm, Federated Averaging (FedAvg), initially proposed by McMahan et al [2], offers a promising solution by enabling multiple clients to collaboratively train a shared global model without exchanging raw data. To be more specific, each client only needs to upload parameters or gradients, which dramatically increases privacy preservation while reducing the need for centralized data storage. 4 years later, FedProx was proposed by Li et al. to handle data and system heterogeneity in FL [3]. The core method of FedProx is to add a proximal term to the local training objective, which makes training easier to converge and helps resolve model drift. In recent years, personalized FL approaches such as FedPlus have emerged, enhancing variants of the FL framework by improving optimization, thereby promoting robustness against malicious clients and stragglers.

Over the years, many studies have explored the application of FL in air quality forecasting and monitoring. To address the air quality forecasting and monitoring challenges, Nguyen, D. V. et al. proposes a spatially distributed federated learning framework combined with CNN recurrent neural networks for air pollution prediction [4]. This manner enables training across multiple geographically dispersed nodes without centralizing raw data. Thus, it can effectively capture the dynamic evolution of pollution in time and space. Based on the problem of data heterogeneity and missing data, Zhou,

X. et al. presents a (FedCGAN) for imputing missing air quality data [5]. By learning the spatial and temporal distribution features of local data. They optimize the generator and discriminator on a global scale, achieving high-quality filling of the void data and robustness. Ramadan et al. combine FL and IoT technology [6], then come up with the HAFA aggregation algorithm for indoor air quality monitoring. They are validating the feasibility of FL in privacy and sensitive building environments. Although there are many advances, the FL in the field of air quality is still facing serious challenges. For example, the challenge of balancing global model convergence with local personalization under heterogeneous, Non-IID, and sparse air quality data, the absence of unified architectural frameworks that effectively integrate FL with edge computing, and deploying in a real scenario to demonstrate the feasibility.

Thanks to the progress and breakthroughs made by researchers in this field. Therefore, it's necessary to make a comprehensive overview. The rest of this passage is as follows. Section 2 (Method) introduced the basic algorithms, such as FedAvg and FedProx . Besides, the passage integrates the probabilistic graphical model, Time-series deep learning methods, Spatio-temporal graph neural network methods, and Multi-model ensemble and transfer learning methods, achieving cross-region collaborative prediction.

Section 3 (Discussion) evaluates the effectiveness of these methods and points out the challenges and insufficiencies. For instance, models lack interpretability and generalization ability when dealing with non-IID data. Meanwhile, this article proposed the direction of development in the future, which includes Cross-domain causal generalization, Improvements to the Adaptive Structure in a dynamic background, and a lightweight heterogeneous federated architecture combined with edge intelligence collaboration. Section 4 (Conclude) summarizes the whole study and evaluates the feasibility of applying FL in air prediction and monitoring. Finally, this paper will integrate the literature reviews to propose the directions and trends for future development in the field.

2. Method

2.1. Preliminaries of federated learning

Federated learning is a distributed machine learning method which permits multiple clients, such as cities, monitoring stations, and IoT devices, to collaboratively train a global model without sharing their private data. Each client trains the model on their local data and periodically sends updates to the central server for aggregation. This approach can provide excellent privacy protection and process the data with relatively high accuracy, which is an important way to monitor air quality.

The most common method is Federated Averaging (FedAvg). In FedAvg, the central server initializes a global model and selects a subset of clients. As described by McMahan et al. [2], "In each round, a subset of clients is selected to train the model locally on their own data for several epochs and then send their updated models to the server, which averages them to produce a new global model." It allows updates many times before aggregation, which prominently reduces the communication costs and protects the data privacy. Although FedAvg is widely used, it will face a challenge when encounter the Non-IID data. The diversity of data may cause the convergence to be slower or lower in accuracy.

Regarding the heterogeneity issue of data distribution for different clients, FedProx and personalized FL were developed to address the problem. FedProx introduces a regularization term on top of the FedAvg to limit the deviation between the local model and global model. It improves the stability and convergence of the heterogeneous data.

Moreover, FL can be combined with deep learning models and graph neural networks to capture the temporal patterns and spatial dependencies of air quality data. FL provides an effective framework that achieves accurate and preserves the privacy of air quality monitoring and prediction.

In this passage, existing models are divided into four groups based on the problems solved by FL in air quality forecasting and the corresponding technical approaches. The classification reflects the evolution from causal explanation to precise prediction, and covers the entire pipeline from single-time-series modeling to complex spatial-temporal modeling, and finally to heterogeneous adaptation in practical deployment.

2.2. Federated Learning Methods and Analysis of Their Application Performance

2.2.1. Probabilistic graphical model methods

Federated Bayesian Networks (FBNs) decentralize the structure learning process, allowing each participant to learn local network structures based on their own data and upload only structural parameters or statistical information to a central server for global integration. This approach builds a cross-region joint causal model, simultaneously maintaining data privacy. Bian and Huang were the first to combine a federated learning framework with Bayesian networks and apply it to cross-regional air quality classification in the Beijing-Tianjin-Hebei region [7]. These three cities are FL clients achieving cross-regional collaborative modeling through horizontal federated learning. To address the issue of getting stuck in local optima, they introduce an adaptive penalty term into the IID data. They use a Markov blanket to break the dependency structure that leads to local optima. When the confidence intervals differ too widely, the system uses weighting to balance the overall weights. Then they propose INIA, which, by simulating the characteristics of probabilistic jumps and the natural selection mechanism of biological immunity, can escape from local optima in Bayesian networks and accurately identify the optimal network structure suited for cross-regional pollution. Besides, using homomorphic encryption technology, clients in each region upload parameters in ciphertext form, allowing the server to perform aggregation without decrypting them, which protects the monitoring data privacy. FBN improved 15%-20% accuracy in the air quality classification and found a strong correlation between NO_2 concentrations and fluctuations of 30% in the AQI index, which provided quantitative evidence.

2.2.2. Time-series deep learning methods

A federated LSTM is the most straightforward implementation approach, in which each client maintains its own LSTM model, and the server performs global aggregation using FedAvg or a variant of it. Ramadan et al. offer the SecureIoT-FL framework, which integrates FL and multi-sensor [8]. Ramadan et al. Presents N-FedAvg as an enhanced version of FedAvg, which uses normalization to balance the data deviation. This paper also combines LSTM with FL, which can precisely capture the long-term dependence. The model will automatically forget the useless data and noise, integrating with the N-FedAvg. In the final evaluation, LSTM demonstrated a better fit for the temporal patterns of the task compared to CNN, which achieved higher accuracy (92% vs 89%) in monitoring the 12 kinds of pollution.

In order to further improve accuracy, Roy et al. presents combining CNN and LSTM [9]. CNN is responsible for extracting spatial local features between monitoring stations, and LSTM is responsible for identifying long-term dependencies in time series data. The research introduces the GAN data completion, which sets the alpha=100(Hyperparameter weight coefficients in the generator loss), hint rate=0.9(Determines how much source information is provided), learning rate=0.001, batch size=256. The article finally indicates that the research has evaluated 26 clients in the city. After analyzing model robustness and generalization ability, the author gets RSME=19.83, RSE=14.41, R^2=0.72, MSE=499.16.

2.2.3. Spatio-temporal graph neural network methods

Liu et al. proposed the GC-LSTM (Graph Convolutional LSTM) model [10], which integrates the GCN and LSTM specifically used to capture the air quality complexity of Spatio-temporal correlation. GCN is responsible for dealing with the spatial topological constructed by the monitoring station, which aims to extract spatial dependency between sites. LSTM takes charge of analyzing historical observation data to catch time dynamics. This dual structure makes the model effectively construct the diffusion and evolution laws of air pollutants in time and space. The author also put forward a lightweight dense-mobilenet, which can convey the picture taken by the drones effectively to different clients. Achieved a balance between prediction accuracy and energy consumption. Besides, it realized high-resolution 3D air quality monitoring. The introduction of federated learning addresses the privacy concerns associated with the inability to share data between ground stations and drones across regions, while also reducing the bandwidth requirements for drone data transmission.

2.2.4. Multi-model ensemble and transfer learning methods

Multi-model federated learning trains multiple base models (such as LSTM, CNN, GRU, etc.) simultaneously within a federated framework, using model fusion strategies to enhance the robustness and generalization capabilities of predictions. Dong et al reviewed the Multi-model FL in the air quality prediction, pointing out that Multi-model fusion can effectively reduce the prediction bias of a single model when data distributions change or model assumptions do not match [11]. So, the author proposed that FL frameworks inherently support parallel training of multiple models, allowing each client to select the optimal model based on the characteristics of its local data. The article indicates that they use different models, such as LSTM, GRU, and TCN, for different clients. The way of Muti-model breaks the traditional manner, which requires the use of the same model for each client. Muti-model FL can choose the suitable model according to each client's computer power. Achieving true heterogeneous compatibility and personalized optimization while ensuring data privacy.

Because the traditional air quality prediction relies on the underground monitoring station. Therefore, Sewtha et al. provide an approach that converts air quality forecasting into a satellite image classification task and uses federated learning to address the issue of data silos [12]. Input the satellite images and output the air quality rank. Use CNN weights pre-trained on ImageNet as initial parameters, then slightly adjust the pre-trained weights. During training, the target-domain client not only fits the local ground truth labels but also imitates the output distribution of the source-domain teacher model. That's why transfer learning enables the rapid development of high-quality predictive models using minimal local data, addressing the challenges of implementing federated learning in cold-start and sparse-label scenarios.

3. Discussion

Methods in Section 2.2, such as Probabilistic graphical model, Time-series deep learning, Spatio-temporal graph neural network, and Multi-model ensemble and transfer learning, have made breakthroughs in each aspect of federated learning applications in air quality monitoring and prediction. However, there are some challenges that impede the study in this area.

3.1. Challenges

3.1.1. Inadequate alignment between model structure and data heterogeneity

Probabilistic graphical models (FBNs) have great performance under the IID assumption, but the performance under a non-IID situation is not that good. The air quality data displays unstable and Non-IID in cross region background. Although FBN introduces the INIA algorithm to escape from the local optimum, robustness against highly heterogeneous or sparse data has not been fully verified.

Muti-model FL permits different clients according to their data characteristics and computer power to select heterogeneous base models (LSTM, GRU, TCN). This format breaks the limitation of traditional FL that each client must use the same type of model. Yet, how to optimally match the optimal model architecture to each client remains an open problem. Besides, convergence is also an issue in terms of theoretical guarantees. A new aggregation algorithm needs to be designed for multi-model environments with convergence guarantees provided.

3.1.2. Model interpretability

In the LSTM, CNN-LSTM model, the features of non-linear internal circulation and control mechanisms make the prediction process a"black box."Decision-makers find it difficult to understand the specific time points or sensor features on which the model's judgments are based, and they are unable to quantify the extent to which each input variable contributes to the final output.

The GT-LSTM has a high accuracy in the short term, but the precision drops dramatically when the duration is long. Nevertheless, the model hasn't explained why the long-term predictions failed, nor is it possible to determine whether the failure was due to a breakdown in spatial topological relationships or to cumulative errors resulting from temporal dependencies.

3.1.3. Versatility of the models

Because some methods introduced are used in a particular region. Means it may not be suitable for other region where have a large climate difference from the trial area. Ramadan et al.'s Federated LSTM achieved a monitoring accuracy of 92% for 12 pollutants [8], but this experiment was conducted under specific sensor network and synchronous aggregation conditions. When faced with inconsistent sensor sampling rates, changes in missing data patterns, or variations in client computing power, can the normalization strategy of N-FedAvg maintain the same level of robustness? Roy et al. CNN-LSTM combined with GAN completion; its hyperparameters (alpha=100, hint rate=0.9) were fine-tuned for specific missing patterns [7]. When the missing patterns are uncertain, the model's performance may drop sharply. Models reflect the current method's limited ability to generalize to data distribution shifts and diverse patterns of missing values.

3.2. Prospects

3.2.1. Cross-domain causal generalization

There are significant differences in the driving factors of pollution across different regions, and causal structures learned from a single region are difficult to directly generalize to other regions with vastly different climatic conditions and pollution source structures. In the future, meta-learning could be incorporated to enable models to rapidly adapt to the data distributions of new regions. It can reduce the impact caused by geographical Variations.

3.2.2. Improvements to the Adaptive Structure in a dynamic background

Due to the evolution of pollution dispersion patterns and changes in meteorological conditions, the result generated by the model is unstable. Developing a dynamic adaptive federated learning framework for air quality monitoring enables the model to update the graph topology during the meteorological changes. According to the situation of data loss, the pattern dynamic adjust the strategy for imputing the missing data. Meanwhile, match the framework with the client side model architectures to local data characteristics automatically. Thus, augmenting the model's robustness and generalization capabilities across different regions, time periods, and sensor conditions.

3.2.3. Lightweight heterogeneous federated architecture combined with edge intelligence collaboration

To tackle the problems of high communication overhead and strong client heterogeneity, future work should concentrate on better the deep integration of lightweight federated learning architectures with edge computing. The SecureIoT-FL framework proposed by Ramadan et al. has already demonstrated that the multi-sensor federated learning is feasible [8]. This approach allows different clients to select heterogeneous base models based on their own computational capabilities. Model pruning and knowledge distillation techniques can be further introduced in the future to generate lightweight student models for resource-constrained edge nodes simultaneously maintaining the high accuracy of the cloud-based teacher model. Based on the feasibility demonstrated by the HAFA aggregation algorithm, which was proposed by Ramadan et al. [6]. In indoor air quality monitoring, a hierarchical asynchronous federated learning framework could be developed in the future. Edge nodes would be responsible for local spatio-temporal feature extraction and preliminary aggregation. The cloud would handle only cross-regional global knowledge fusion. This approach would enable the real-time deployment of large-scale, multi-city federated air quality monitoring systems and ensuring privacy.

4. Conclusion

This paper finishes a comprehensive review of the applications of federated learning in air quality forecasting and monitoring. Following the investigation, federated learning is a promising approach which enables collaborative model training across distributed clients without sharing raw data. It integrates probabilistic graphical models, time-series deep learning, spatio-temporal graph neural networks, and multi-model ensemble strategies. After analyzing these methods, it is shown that they improve prediction accuracy and capture temporal-spatial dependencies. Personalized FL approaches contribute to data heterogeneity and improve convergence. However, there are some challenges that remain, such as struggling with Non-IID data, limited interpretability, and generalization across regions with diverse climates and pollution sources. That's why future research should focus on developing adaptive federated learning models that can generalize across regions and dynamically adjust to changing meteorological conditions and pollution patterns. Overall, federated learning provides a feasible and scalable framework for privacy-preserving, high-resolution air quality monitoring and forecasting. Making significant contributions to the field of research and practical deployment.

References

[1] Zhang XX, et al. Aeolian dust in Central Asia: Spatial distribution and temporal variability. Atmos Environ. 2020;238:117734.

[2] McMahan B, et al. Communication-efficient learning of deep networks from decentralized data. In: Proc Artif Intell Stat (AISTATS); 2017.

[3] Li T, et al. Federated optimization in heterogeneous networks. Proc Mach Learn Syst. 2020;2:429–450.

[4] Nguyen DV, Zettsu K. Spatially-distributed federated learning of convolutional recurrent neural networks for air pollution prediction. In: Proc IEEE Int Conf Big Data; 2021.

[5] Zhou X, et al. Federated conditional generative adversarial nets imputation method for air quality missing data. Knowl Based Syst. 2021;228:107261.

[6] Ramadan MNA, Ali MAH, Alkhedher M. Development of a federated learning-enabled IoT framework for indoor air quality and HVAC optimization in healthcare buildings. J Build Eng. 2025;107:112758.

[7] Bian C, Huang G. Federated Bayesian network approach for cross-regional air pollution classification: A case study of the Beijing–Tianjin–Hebei region. Environ Monit Assess. 2024;196(7):668.

[8] Ramadan MNA, et al. SecureIoT-FL: A federated learning framework for privacy-preserving real-time environmental monitoring in industrial IoT applications. Alex Eng J. 2025;114:681–701.

[9] Roy S, Kandhoul N, Dhurandher SK. Generative adversarial network-based imputation and personalized federated learning with convolutional neural network-long short-term memory for air quality index prediction. Eng Appl Artif Intell. 2025;158:111300.

[10] Liu Y, et al. Federated learning in the sky: Aerial-ground air quality sensing framework with UAV swarms. IEEE Internet Things J. 2020;8(12):9827–9837.

[11] Le DD, et al. Insights into multi-model federated learning: An advanced approach for air quality index forecasting. Algorithms. 2022;15(11):434.

[12] Swetha G, Datla R, Mohan CK. Federated learning to categorize air quality using satellite images. IEICE Proc Ser. 2024;81:57–60.