PmForecast: Leveraging Temporal LSTM to Deliver In situ Air Quality Predictions

Maryam Rahmani, Suzanne Crumeyrolle, Nadège Allegri-Martiny, Amir Taherkordi, Romain Rouvoy

PmForecast: Leveraging Temporal LSTM to Deliver *In situ* Air Quality Predictions

Maryam Rahmani[1*], Suzanne Crumeyrolle[2], Nadège Allegri-Martiny[3], Amir Taherkordi[4], Romain Rouvoy[1]

[1*]Univ. Lille, Inria, CNRS, UMR 9189 CRIStAL, France.
[2]Univ. Lille, CNRS, UMR 8518 LOA, France.
[3]Univ. Bourgogne, CNRS, UMR 6282 Biogéosciences, France.
[4]Univ. Oslo, IFI, Norway.

*Corresponding author(s). E-mail(s): maryam.rahmani@inria.fr;
Contributing authors: suzanne.crumeyrolle@univ-lille.fr; Nadege.Allegri-Martiny@u-bourgogne.fr; amirhost@ifi.uio.no; romain.rouvoy@univ-lille.fr;

Abstract

The physical and chemical properties of atmospheric aerosol particles are crucial in influencing global climate and ecosystem processes. Given the numerous studies highlighting adverse health effects from exposure to aerosol particulates, particularly PM, effective air quality management strategies are under consideration (Annesi-Maesano et al, 2007). Herein, we introduce a predictive model—PmForecast—employing a self-adaptive LSTM architecture to predict $PM_{2.5}$ values in the real atmosphere. Specifically, we explore adopting a T-LSTM model to better benefit from temporal dimensions. PmForecast is strategically designed with four key phases: preprocessing, temporal attention, prediction horizon, and LSTM layers. By leveraging LSTM's significant predictive ability in time-series data, the inclusion of temporal attention enhances the model's specificity. Temporal dynamics modeling entails generating insights over time, utilizing temporal attention to extract essential characteristics from historical air pollutant concentrations, with the flexibility to adjust the historical data according to the forecasting period. To assess PmForecast, we consider measurements collected from the QameleO network, a sparse network of air-quality micro-stations deployed in Dijon, France. The self-adaptive capabilities of PmForecast allow the model to be dynamically updated, evaluating its performance and continuously tuning hyper-parameters based on the latest data trends. Our empirical evaluation reports that PmForecast outperforms the state of the art, achieving notable accuracy in both short-term and long-term predictions. The PmForecast deployment at scale can serve as a valuable tool for proactive decision-making and targeted interventions to mitigate the health risks associated with air pollution.

Keywords: Air Quality Forecasting, Air pollution, PM2.5 Prediction, Urban Air Quality, LSTM

1 Introduction

Cities and their citizens face critical problems that result from the growth of industry and urban regions. As highlighted by the *World Health Organization* (WHO) World Health Organization (2006), this growth represents one of the most significant environmental health risks and remains a leading cause of mortality. Specifically, air pollution stands out as a primary contributor to various health issues, including lung cancer and bronchial asthma (Cohen et al, 2017; Mikati et al, 2018; Philip et al, 2017; Harrison, 2016; Hu et al, 2023). In addition, air pollution adversely affects both terrestrial and aquatic ecosystems, leading to environmental degradation and loss of biodiversity. This highlights the importance of addressing air pollution to mitigate environmental issues effectively (Manisalidis et al, 2020).

Particulate Matter (PM) is a key component of air pollution found in the atmosphere and classified by size into PM_{10}, $PM_{2.5}$, and PM_1 categories (based on micrometer size). PM is known to significantly impact the cardiovascular and respiratory systems, leading to a spectrum of health problems (World Health Organization, 2006). The finer fraction of PM ($PM_{2.5}$ and especially PM_1) can penetrate deeply into the lung, posting detrimental health effects (prabhu and Shridhar, 2019). Moreover, $PM_{2.5}$, observed in urban areas (Guo et al, 2024), is enriched in hazardous metals and organic compounds (Zhang et al, 2020), potentially inducing additional oxidative stress (Terzano et al, 2010). This

pollution not only endangers immediate health but also jeopardizes the long-term sustainability of urban living environments.

Over the past decades, significant research efforts have been directed toward air quality forecasting, especially $PM_{2.5}$. These studies can be clustered into two distinct groups. On the one hand, the first group relies on physical methods, utilizing numerical simulation models based, among other things, on emission inventories, meteorological fields, physical and chemical processes for pollutant transport, dispersion, pollutant aging, and data assimilation (satellite, *in situ* and remote sensing networks, etc.) to forecast air quality (Lee et al, 2017; Ponomarev et al, 2020; Powers et al, 2017). For instance, Powers *et al.* (Powers et al, 2017) employed the *Weather Research and Forecasting* (WRF) model, a widely-used numerical weather prediction model applied globally with convection-permitting resolutions such as 3-km grids. While the WRF Model has demonstrated advancements in operational forecasting and fine-scale atmospheric simulation, additional challenges in physical models for air quality forecasting warrant further exploration. These challenges encompass accurately representing complex atmospheric processes, integrating real-time observational data, addressing uncertainties in emission inventories, and optimizing model parameters for improved predictive capabilities. Investigating these limitations is essential to enhance the accuracy and reliability of air quality forecasts. Furthermore, physical models face constraints related to the quality of pollutant inventory emissions, intricate calculations, and the elevated uncertainty of forecasted results (Wang et al, 2019). On the other hand, the second group leverages statistical algorithms to predict air pollutants from a time series analysis perspective. Statistical predictive methods are based on a modeling approach to predict upcoming air quality, relying on historical $PM_{2.5}$ time series. In contrast to physical forecast methods, statistical approaches involve less computation, as they circumvent the intricate mechanisms of pollutant emission, diffusion, aging, and deposit processes.

A diverse range of statistical methods, encompassing both linear and nonlinear algorithms, has found widespread application. Common approaches include the utilization of *Machine Learning* (ML) algorithms, such as *Support Vector Regression* (SVR) (Zhu and Hu, 2019) and *Autoregressive Integrated Moving Average* (ARIMA) (Che Lah et al, 2023), which are notable for their linear characteristics. For instance, Bassirou Ngom *et al.* (Ngom et al, 2021) present a unique integration of system observations from various stations with a multi-agent simulation, providing a model for assimilating PM_{10} pollution data through a real-time simulation based on the autoregressive ARIMA method. Additionally, *Gaussian Process Regression* (GPR) (He et al, 2023), *gradient boosting* (XGBoost) (Li et al, 2022), *Artificial Neural Network* (ANN) (Guo et al, 2023b; He et al, 2022; Guo et al, 2023a) and deep learning algorithms, like *Recurrent Neural Networkss* (RNNs) (Tu and Wu, 2022) offer effective solutions, leveraging the power of non-linearity to capture complex (or non-linear) relationships in diverse datasets. Some studies, such as (Guo et al, 2020, 2023c), show that combining correlation analysis with ANNs and wavelet-enhanced ANNs (WANNs) effectively reveals both linear and nonlinear relationships between air pollution indices (API) and meteorological variables.

Among these, RNN methods have attracted considerable attention from researchers due to their capacity to improve the correlation between input and output data and the time-dependent nature of the data. Several scientific studies highlight the exceptional ability of RNN-based models to capture temporal dependencies within $PM_{2.5}$ input data, making them effective tools to predict $PM_{2.5}$. Significant research efforts have been directed toward air quality forecasting, utilizing *Gated Recurrent Unit* (GRU) (Zhang et al, 2022b; Panneerselvam and Thiagarajan, 2023) and particularly LSTM (Alhirmizy and Qader, 2019; Gangwar et al, 2023; Zaree and Honarvar, 2018; Bui et al, 2018). Zhang Qi *et al.* (Zhang et al, 2022a) stand out by integrating domain-specific features and a hybrid CNN-LSTM structure, achieving superior accuracy in fine-grained air pollution estimation and predicting compared to compatible baselines. Verma Ishan *et al.* (Verma et al, 2018) introduce a bidirectional LSTM model to predict $PM_{2.5}$ severity levels. The approach significantly improves prediction accuracy by leveraging an ensemble of 3 bidirectional LSTM and incorporating weather data across multiple locations in New Delhi. Yi-Ting Tsai *et al.* (Tsai et al, 2018) introduce an approach utilizing RNN with LSTM using US *Environmental Protection Agency* (EPA) data, effectively predicting $PM_{2.5}$ values for the next 4 hours at 66 monitoring stations in Taiwan with promising results. (Erbiao and Guangfei, 2023) proposes an innovative hybrid self-attention model LSTM, which bridges the gap by efficiently predicting long-term $PM_{2.5}$ concentrations in classrooms, outperforming existing methods, and significantly reducing computational complexities. Moreover, Abimannan *et al.* (Abimannan et al, 2020) propose a model based on LSTM / Multivariate Variate Regression -MVR- to improve $PM_{2.5}$ prediction accuracy, particularly during summer and winter. Comparative analysis with LSTM reveals that the proposed LSTM/MVR model's efficiency predict hourly $PM_{2.5}$ with higher precision.

Despite the notable progress in air quality predictions using LSTM-based models, the field faces persistent challenges and limitations (Srivastava and Kumar Das (2023)). In particular, the accuracy of predictions hinges heavily on the quality and representativeness of sequential data, with incomplete or biased datasets, potentially compromising model performance. Additionally, while LSTM models excel in capturing temporal dependencies, they may encounter difficulties with abrupt data changes or outliers, necessitating further refinement for robust predictions. Finally, deploying such models at scale requires careful considerations for computational efficiency and online

processing, especially in urban areas where a significant volume of data could be generated and collected.

In response to the challenges described above, we present our Temporal LSTM forecasting model (PmForecast), designed specifically to address the complexities of urban air quality prediction. In addition to leveraging the strengths of standard LSTM architecture, our model adeptly captures temporal dependencies and integrates mechanisms to address data quality issues, thus enhancing predicting accuracy. Furthermore, the embedded temporal mechanism in our model contributes to robust and sustainable long-term predictions. A notable aspect is our utilization of locally available and cost-effective sensors from existing devices. This approach not only enriches the model's adaptability but also enhances its accessibility. Through meticulous optimization of the PmForecast model, we aim to surpass traditional simulation methods, providing a resource-efficient alternative that reduces both time and energy consumption, ultimately establishing a local real-time framework. Extensive testing and validation have demonstrated the high accuracy of our PmForecast model in long-term forecasting, outperforming existing models and contributing to the advancement of sustainable and precise air quality predictions for urban communities.

The structure of the manuscript sections is described as follows: Section 2 delves into a detailed explanation of the methodology employed. Section 3 outlines the instruments used and the associated data. Following that, Section 4 presents and discusses the experimental outcomes derived from our proposed model. We conclude and suggest potential avenues for future research in Section 5.

2 Forecasting Particulate Matter with PMForecast

As we embark on the pivotal task of predicting $PM_{2.5}$ concentrations for both immediate and future time frames in urban environments, the foundation of our approach is rooted in the strategic choice of a neural network architecture. Our conviction in the relevance of the LSTM model stems from its exceptional ability to discern and interpret temporal patterns, a crucial aspect in unraveling the intricate dynamics of air quality. By harnessing the power of LSTM architecture, our model, PmForecast, stands as a testament to our commitment to precision in predicting.

Figure 1 depicts an overview of the PmForecast framework developed in this study. The proposed model consists of four primary phases: preprocessing, temporal attention, prediction horizon, and LSTM layers. The network is fueled by historical observations, and the outputs encompass the temporal dynamics of the predicted values. Observations are recorded by the sensor every 15 minutes, amounting to four times per hour. To reduce granularity, we resample the data to one point per hour, given the expectation of minimal change within a 60-minute time frame. Additionally, the model can operate online by evaluating and retraining at specified intervals, contingent upon available data.

2.1 LSTM Model

Our goal centers on optimizing ML algorithms for superior performance, with a specific focus on deep learning techniques, particularly RNNs, which have demonstrated effectiveness in processing sequential data. However, traditional RNNs face challenges in long-term prediction tasks, prominently contending with issues such as gradient disappearance (Noh, 2021). This phenomenon occurs during the RNNs training process, when the loss function gradients concerning network parameters diminish significantly as they are back-propagated through time.

To address this challenge, advanced variants like LSTM models introduce mechanisms, including *gating*, to enhance the accuracy of long-term predictions—an essential consideration in our work specifically focused on air quality forecasting. In an LSTM cell, three gates—the input, forget, and output gates—play a crucial role in mitigating the challenges posed by gradient disappearance (Le and Zuidema, 2016). For a comprehensive understanding of the architecture and functioning of LSTM models, we recommend referring to the seminal work by Hochreiter and Schmidhuber (Hochreiter and Schmidhuber, 1997). Leveraging the LSTM model as the foundational model, we aim to achieve optimal performance among various machine learning algorithms, specifically including GRU, GPR, XGBoost, and ARIMA.

2.2 Temporal Dynamics Modeling

Afterward, the network structure needs to be optimized in terms of time, cost, and performance. To achieve this, we aim to leverage *Time-Focused Insight Generation*. Inherently considering temporal correlations of historical air pollutant data helps to improve performance.

Temporal Attention

We proficiently capture the essential characteristics from historical environmental data through temporal windows whose duration is dynamically adjusted depending on how far ahead the prediction is being made.

Subsequently, we utilize the LSTM layer to extract temporal information from these mapped features. In time series, incorporating time-variant features is pivotal for capturing effective temporal dynamics. We systematically extract supplementary time-related features due to their strong influence on PM predictions such as (i) weekdays versus weekends related to traffic and industrial emissions and (ii) hours of a day related to hourly emission strength modification and the rise of the boundary layer height (i.e. the lower part of the atmosphere influenced by the Earth's surface) influencing the dilution of particles in the atmosphere over the course of the day.

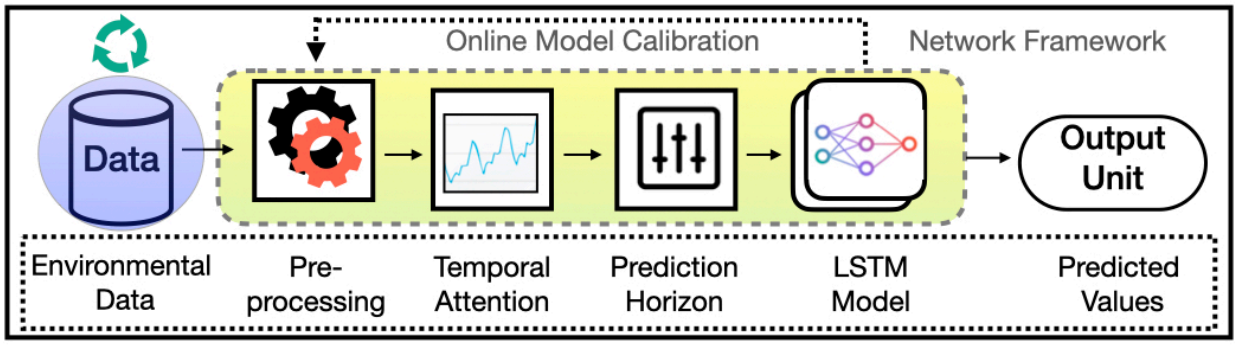


Fig. 1: The comprehensive framework of PmForecast designed for air pollution prediction is outlined, comprising four key steps: data pre-processing, temporal attention to mitigate gradient disappearance, a flexible prediction horizon for dynamic future forecasting, and layers employing Long Short-Term Memory (LSTM)—the trainable component. Further details are provided in Section 2.1. The term 'Environmental data' pertains to data previously collected and utilized by the model for training purposes.

Recognizing the significance of these temporal factors, we seamlessly integrate them as essential features in our model's input by merging them with the measurements. When optimizing temporal history for prediction, one question arises: How much of the historical environmental data should be considered?

An extensive sensitivity analysis, in which different lag times were tested, reveals that the lag time should be dynamically adjusted based on the prediction's time horizon to obtain the most accurate PM predictions. For one-hour-ahead forecasting, our experimental results indicate that setting a$lag_{constant}$ at 3 hours consistently produces optimal outcomes. This choice was derived through extensive experimentation, where we tested various lag times, training our model with different historical data windows $[1, 3, 6, 9, 12]$hours, and evaluating their prediction accuracy using metrics like *Coefficient of determination* (R^2) and *Root Mean Square Error* (RMSE). The 3-hour window consistently produced the lowest prediction errors, indicating optimal performance for capturing relevant temporal patterns. Further validation confirmed that adding more historical data did not significantly improve accuracy, and shorter windows led to higher errors, establishing the 3-hour window as the best balance for accurate predictions. For extended prediction horizons, the same experiments were conducted. As the prediction time extends, the lag time incrementally increases, following the empirical relationship we derived, encapsulated in Formula 1. Specifically, for every 6-hour extension in the prediction horizon, the increment in the lag_t results in the incorporation of additional past observations, enhancing performance.

$$\mathrm{lag}_t = \mathrm{lag}_{\mathrm{constant}} + \mathrm{round}\left(\frac{\mathrm{pre}_t}{\mathrm{lag}_{\mathrm{rate}}}\right) \quad (1)$$

Equation 1 is derived from our experimental findings, which suggest a dynamic relationship between the prediction horizon and the optimal lag time. It is crucial to highlight that, in our experiments, pre_t denotes the prediction period and the constant lag_{rate} is set at 6, representing increments based on predictions made every 6 hours.

Prediction Horizon Strategies

We formulate a strategy tailored to meet long-term prediction demands. Subsequently, our framework excels at forecasting air pollution intervals based on user preferences, specifically for the next few days with a time stamp interval of 1 hour. The mechanism dynamically updates the ground truth data, lag time observations, and the output unit according to the user's preferences. This process is visually represented as the *Prediction Horizon* in Figure 1. When the user modifies preferences, online updates reconfigure the pre-processing and dynamics of the temporal attention mechanism to align with the new purpose. Then, the model is retrained.

3 Sensor Deployment

This section provides detailed information of the observations used to feed our model (QameleO network in Dijon), including their acquisition and pre-processing.

3.1 The QameleO Network

QameleO is a *inexpensive* air quality micro-station developed by two research teams in the University of Burgundy and *Institut de Recherches pour le Développement* (IRD) (Martiny et al, 2023). QameleO microstations measure the mass concentrations within PM_1, $PM_{2.5}$, PM_{10} fractions along with meteorological

variables, such as temperature and relative humidity. The measurements are consistently available every 15 minutes, aligning with the time-step of the stations of the approved air quality monitoring association (AASQA : Association agréée de surveillance de la qualité de l'air) and operated by regional governments in the Dijon Metropolis.

QAMELEO micro-stations have supported tests in the laboratory and outdoors, in the frame of a national evaluation exercise leaded by the LCSQA (Central Laboratory of Air Quality Surveillance) in July 2018. This proficiency testing of micro-sensors systems, referred to as the $EA\mu C$ field campaign, enabled to compare the QameleO micro-station to 15 other micro-sensors and to the BAM1020 reference analyser, measuring the PM2.5 fraction, for 2 entire weeks and at the same location/station. These tests established that the QAMELEO micro-station can satisfactorily reproduce the temporal dynamics of the PM mass concentrations (Crunaire et al, 2018; Redon et al, 2018).

In particular, for $PM_{2.5}$, the correlation coefficient between the micro-station and a referenced station is $+0.73$, which is a significant score at a 99% level according to the Bravais-Pearson statistical test, and a mean bias of $-2.71 \mu g/m^3$.

In Dijon Metropolis, the POPSU (*Plateforme d'Observation et de Stratégies Urbaines*) program has been a real opportunity to deploy the QAMELEO network in a real urban environment (cf. Figure 2). The QAMELEO microstations were implemented like meteorological stations, all under the same conditions: at 3 meters high, with a similar Sun Exposition. There are 12 QAMELEO micro-stations implemented in Dijon Metropolis. Of the 12 micro-stations, four of them cover a complete year (from November 2020 to October 2021) of measurements as the network has been deployed in progressive phases. These four specific stations are located within the city in *Port du Canal,* Hoche, *Carnot, and* Janin, representative of diversified urban conditions (traffic : *Carnot and* Hoche, urban background: *Port du Canal and* Janin).

PMFORECAST was used on each station independently to test its efficiency.

3.2 Data Preprocessing

The data preprocessing stage is crucial in improving the quality and suitability of the data set for comprehensive analysis. It involves a meticulous process of cleaning, transformation, and organization to ensure data accuracy and consistency while eliminating errors. QAMELEO dataset is validated and corrected for the concentrations of the PM mass according to the Isolation Forest (IF) method developed by the University of Burgundy. The Isolation Forest algorithm is an unsupervised machine learning technique specifically designed for anomaly detection. It works by isolating observations through a random selection of features and split values, efficiently identifying anomalies due to their distinct nature. (Martiny et al, 2023) employed this method to identify and correct anomalous concentration values of the mass of PM, ensuring the precision and reliability of the data set that we used in our study. Addressing missing values is a crucial step before diving into data analysis. Despite the QAMELEO microstations offering a relatively consistent dataset for air quality assessment, with an average of approximately %5 missing values over a year time series, ML methods require a dataset without gaps. To meet this criterion, we applied a 12-hour moving average.

In the final stage, where a few minor missing values persisted around 0.8% to 0.9% percent, we opted for forward filling, replacing each missing value with the most recent observed value in the dataset. Achieving uniformity in the dimension values is crucial for meaningful analysis. To ensure this, we employed *Min-Max Normalization*, a technique that scales the dimension values to a range between 0 and 1. This normalization process contributes to equitable data representation, a fundamental aspect of robust analysis. The final versions of the datasets are prepared for four sites: Station 1 (Canal), Station 2 (Hoche), Station 3 (Carnot), and Station 4 (Janin). We include measurements spanning 9.5 months in the training sets and 2.5 months in the testing sets for all stations.

3.3 Model Hyper-parameters

In our pursuit of creating an optimal configuration for the LSTM model, the fundamental aspect of our goal is to minimize hardware requirements and employ a lightweight model while maximizing performance. The number of hidden units and layers within a neural network are crucial hyper-parameters that significantly impact the model's capacity and complexity. We strongly emphasize achieving a delicate balance between the model's capacity and the potential risks of over-fitting or under-fitting. In this endeavor, we carefully considered the unique characteristics of the problem and dataset, ensuring that our model is not only efficient but also tailored to the specific challenges posed by the given context. Our investigation comprehensively assesses how varying numbers of hidden layers and units impact the model's performance. Conducting an exhaustive analysis, we explored multiple unit configurations within the range of $[32, 64, 128, 256]$, along with variations in the number of layers ranging from 1 to 5. Model evaluation was performed using R^2 and RMSE metrics for four sites in both train and test datasets. Ultimately, we determined that the optimal architecture for our LSTM model comprises 2 hidden layers with 128 units each, utilizing the *Relu* activation function. The model was trained using the mean squared error *Mean Square Error* (MSE) for the loss function, chosen for its effectiveness in regression tasks.

During the training phase, it is crucial to specify hyper-parameters that significantly influence the performance of deep learning models. Firstly, the *learning rate*, a parameter that determines the size of the steps taken during the model optimization process, was set at 10^{-3}, after evaluating values within the

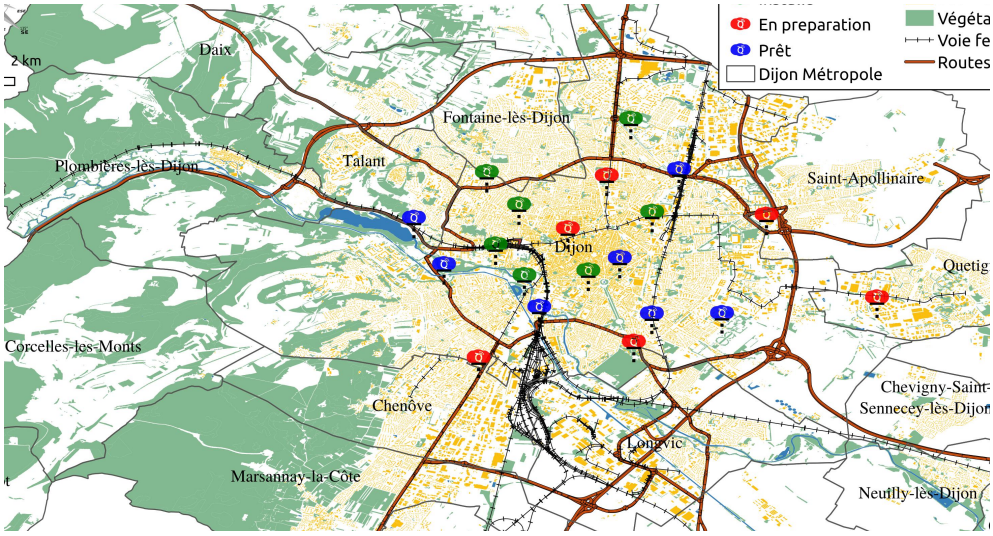


Fig. 2: Locations of Air Pollution Monitoring Micro-Stations in Dijon. The blue circles in the black box correspond to the four QAMELEO stations used in this study (Martiny et al, 2023)

range $[10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}]$. Next is the *batch size*, representing the number of data samples processed in one iteration during model training. A carefully chosen batch size of 48 was implemented, indicating that the model processed 48 (equivalent to 2×24 hours) training examples per iteration. Lastly, the *epoch* parameter, denoting the number of complete passes through the entire dataset during model training, was set to an extensive value of 200. This choice allowed the model to iterate through the entire training datasets 200 times, capturing intricate patterns and enhancing overall performance.

To avoid overfitting and promote model generalization, the *Early Stop* technique was used. This involved monitoring the model's performance on a validation set during the training process and interrupting training once the performance ceased to improve or started to degrade, effectively preventing unnecessary further training.

3.4 Dynamic Datasets & Online Model Calibration

Our model is designed for monthly updates, as we typically observe minimal changes within a one-month timeframe. Embracing a dynamic approach, the model undergoes regular fine-tuning of its hyperparameters based on processed information. It continuously assesses performance metrics, such as accuracy (R^2) and root mean square error (RMSE), selectively incorporating updates when improvements are detected. This iterative process ensures efficient training over the specified timeframe, maintaining the model's currency, and optimizing the latest data trends.

Furthermore, any changes in user preferences can be applied online, allowing the model to be quickly recalibrated to better suit user needs. The adaptability of this approach allows our model to respond effectively to evolving patterns and progressively improve its performance over time.

4 Experimental Results

A thorough series of assessments were conducted to comprehensively present our results. Beginning with an evaluation of hourly prediction precision for the Canal site, we then delve into an in-depth analysis of extending time-frame predictions across all sites. Following this, we undertake a comparative study involving various popular ML algorithms for time-series forecasting. Additionally, we assess the feasibility of multi-task prediction. Finally, we examine the time-consuming aspects across each phase of our framework.

4.1 Precision of Air Pollution Forecasting

One of the main objectives of this research is to achieve high precision in the prediction of air pollution, and in particularly $PM_{2.5}$. The experimental results displayed in Figure 3 provide a visual representation of $PM_{2.5}$ readings over time for the Canal site, illustrating both predicted and ground truth values from the train dataset and test dataset. In a specific time-frame, Figure 4 showcases the ground truth and predicted values for the test set on 24th and 25th July 2021.

Following the assessment of model performance presented in Tables 1 and 2, it is essential to delve into the specific evaluation metrics employed. The assessment includes not only RMSE and R^2, but also incorporates other key performance metrics, such as MSE, *Mean Absolute Error* (MAE), and *Weighted Average of Absolute Percentage Error* (WMAPE). RMSE quantifies the average magnitude of prediction errors, providing a comprehensive measure of model accuracy. MSE offers a similar insight without considering the square root, emphasizing larger errors. MAE represents the average absolute difference between predicted and actual values, offering a robust measure of model precision. R^2 gauges the share of correctly predicted instances, providing a holistic view of model effectiveness. WMAPE, calculated as the weighted average of

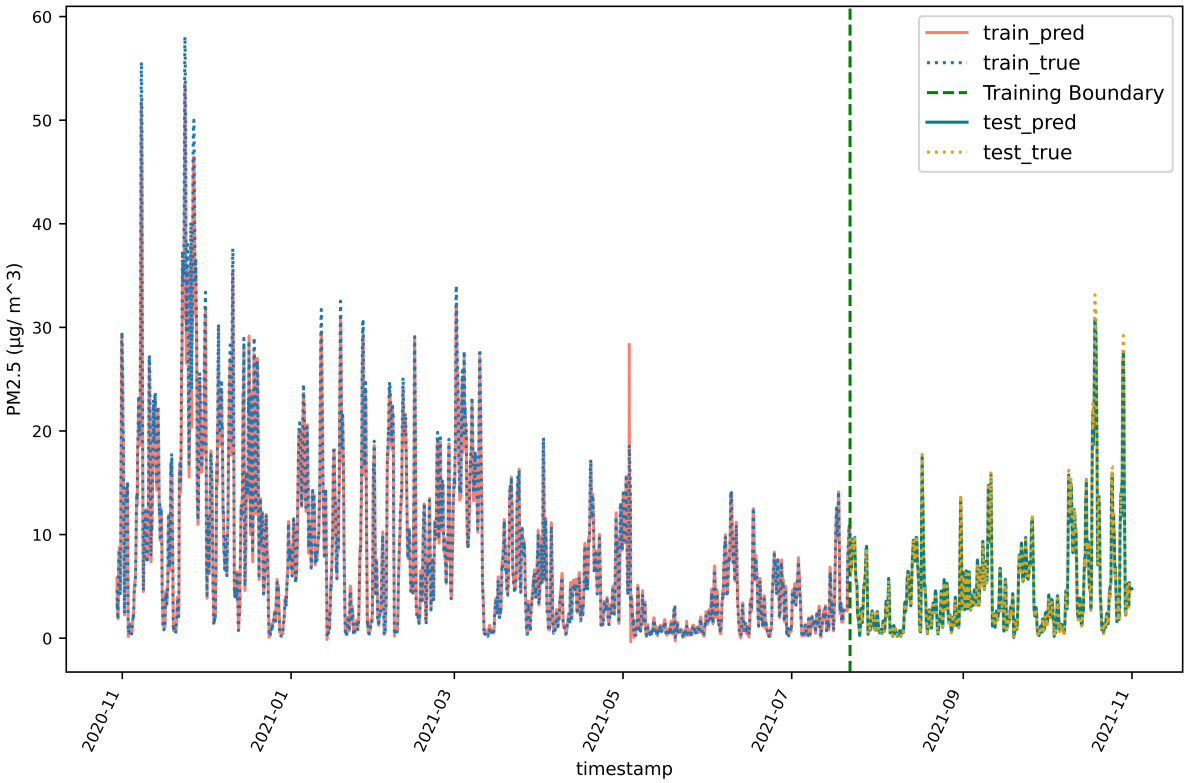


Fig. 3: Hourly temporal prediction of $PM_{2.5}$ levels over time for Canal site. The dotted lines correspond to the observed values and are representative of the true values during the training (blue) and prediction (golden) periods. The solid lines correspond to the $PM_{2.5}$ predicted during the training (salmon) and the prediction (green) periods. The dashed vertical green line indicates the division between the training and test datasets.

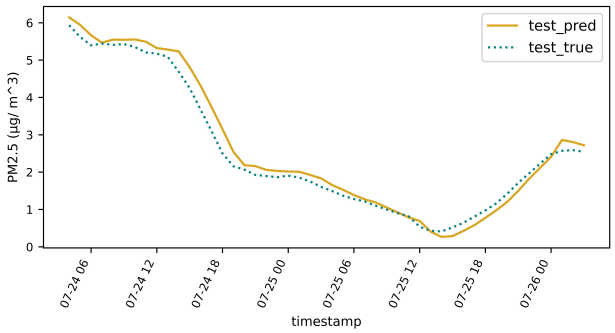


Fig. 4: Hourly temporal prediction of $PM_{2.5}$ levels over time over the Canal site, forecasting 2-day predictions for July 24th (Saturday) and July 25th (Sunday), 2021. The golden solid line represents the predicted values and the dotted green line represents the truth values for the test set.

absolute percentage errors, offers a nuanced perspective by considering the significance of errors across different prediction scenarios. These metrics underscore the model's robust performance across diverse evaluation criteria for training and test datasets.

The predictions show remarkable precision, achieving an impressive R^2 of approximately 100% and substantial RMSE ranging from 0.36 to 0.77 $\mu g/m^3$ in the train and test sets at all stations, respectively. This signifies a robust correlation between predicted and measured $PM_{2.5}$, highlighting the model's exceptional predictive capabilities for the subsequent hour. The combination of high R^2 and low RMSE underscores the reliability and precision of the model in capturing and forecasting target values. Specifically, we note metrics with values less than $0.43\mu g/m^3$, $0.6\mu g/m^3$, and 11% for MAE, MSE, and WMAPE, respectively. In the context of our study, it is essential to acknowledge that our experimental setup involves small test datasets and a model of relative simplicity. In light of these considerations, it is observed that RMSE exhibits a lower value in the test dataset in comparison to the training dataset, which is a general trend observed in machine learning experiments.

While all results demonstrate significance, the notable prominence of the Canal station in the test set and the Carnot station in the train set as the most favorable matches suggesting the best alignment between observed and predicted values at these specific locations (Tables 1 and 2). Bold values in these tables highlight the highest performance across all sites. Despite local attributes such as unique environment leading to different emission sources as well as their diurnal variations, geographical, or meteorological conditions inherent to each station, the model is surprisingly performing well across all sites. This robust adaptability underscores the model's effectiveness across diverse environmental conditions within this city. More metropolises need to be tested to confirm this behavior with probably more contrasted typology (rural vs urban).

Table 1: Evaluation metrics (RMSE, MAE, MSE, R^2, WMAPE) for prediction results during the **training** period for the 4 QAMELEO stations (Canal, Hoche, Carnot, Janin) focusing on 1-hour predictions with a history of 3 hours. **Bold values** indicate the best performance across all sites.

Train-set	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	MSE ($\mu g/m^3$)	R^2 (%)	WMAPE (%)
Canal	0.566	0.352	0.320	**99.5**	**0.043**
Hoche	0.768	0.428	0.590	98.6	0.078
Carnot	**0.438**	**0.262**	**0.192**	99.3	0.054
Janin	0.626	0.414	0.392	99.4	0.047

Table 2: Evaluation metrics (RMSE, MAE,MSE,R^2, WMAPE) for prediction results during the **testing** period for the 4 QAMELEO stations (Canal, Hoche, Carnot, Janin) focusing on 1-hour predictions with a history of 3 hours. **Bold values** indicate the best performance across all sites.

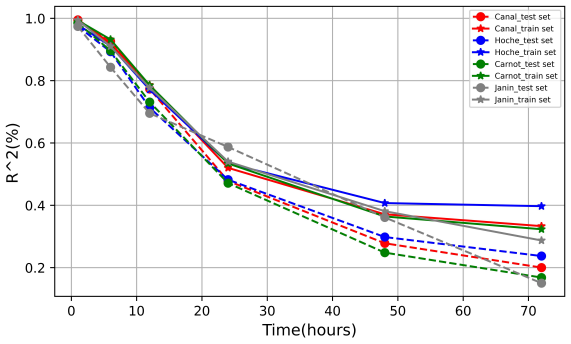
Test-set	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	MSE ($\mu g/m^3$)	R^2 (%)	WMAPE (%)
Canal	**0.353**	0.238	**0.125**	**99.4**	**0.049**
Hoche	0.393	0.241	0.154	98.2	0.106
Carnot	0.357	0.200	0.164	98.7	0.071
Janin	0.444	**0.179**	0.197	97.4	0.061

4.2 Extended Time-frame Prediction

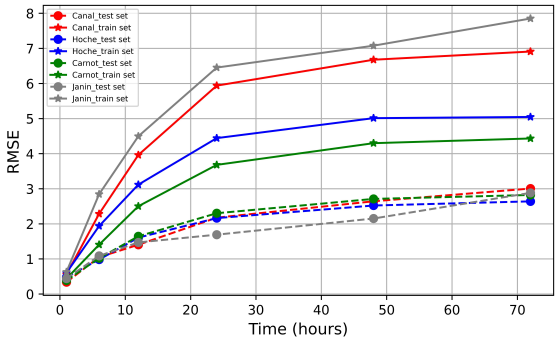
In Figure 5, we assess the accuracy of $PM_{2.5}$ predictions in the train and test sets across datasets from the four stations, ranging from 1 to 72 hours into the future. Figure 5a depicts the computed score, representing the average over the prediction period. Leveraging hourly predictions with adaptable horizons for the near future, we conducted an examination to evaluate performance across various timeframes—specifically $[1, 6, 12, 24, 48, 72]$ hours for all stations. This resulted in $[1, 6, 12, 24, 48, 72]$ hours of prediction values for each time horizon. To provide an example, when forecasting $PM_{2.5}$ levels for the next 12 hours, the model generates a predicted value for each hour within this period, amounting to 12 values for this configuration. The R^2 score for the entire 12-hour prediction horizon is subsequently calculated by averaging these 12 R^2 values.

As the prediction period extends, a noticeable decrease in accuracy for each individual hour is observed, as depicted in Figure 5. Another noteworthy phenomenon emerges with the extension of the time horizon, where the performance decreases for a specific time compared to a shorter future prediction. For instance, with a 24-hour forecasting horizon, the accuracy for the first hour drops to approximately 91.5%, while a 1-hour forecasting horizon achieves a higher accuracy of approximately 99%. In particular, for the 48-hour horizon, R^2 drops to 0.4 for all sites in both the train and test periods, indicating a limitation in the model's ability to predict PM levels beyond 36 hours.

This observed behavior aligns with the common phenomena noted in both physical and numerical models Zhao et al (2023). While the results showcase the model's ability to generalize and provide reliable forecasts for the near future, challenges arise when forecasting for more extended timeframes.



(a)



(b)

Fig. 5: Performance Evaluation of Long-Term $PM_{2.5}$ Forecasting Across Multiple Sites: (a) Accuracy Assessed by R^2 % metrics, and (b) Root Mean Squared Error $RMSE \mu g/m^3$. The solid lines with stars denote the performance on the training sets, while the dashed lines represent the performance on the test sets. Each of the four stations is distinguished by a unique color: Canal (red), Hoche (blue), Carnot (green), and Janin (grey).

4.3 Method Comparison Study

Our study considered a diverse set of widely-recognized time series data forecasting algorithms, including *Gaussian Process Regression* (GPR), *Gated Recurrent Units* (GRU), *XGBoost*, *ARIMA*, and Standard LSTM. We evaluate these ML algorithms to forecast $PM_{2.5}$ levels over 1- and 12-hour periods, carefully examining their performance over time across four different sites. The consistent findings across these sites lead to a comprehensive analysis presented in Table 3, showcasing the effectiveness of each algorithm in meeting our forecasting objectives. GRU demonstrates performance very close to PmForecast, attributed to their similar architectures and shared algorithms in RNN models. However, PmForecast consistently outperforms GRU, displaying superior results. ARIMA excels in the short term, providing accurate predictions, but its computational demands increase for longer forecasting horizons, leading to less efficient performance. XGBoost, despite its stability and rapid training times, falls short compared to RNN-based algorithms. GPR achieves high rankings on training data but delivers less favorable outcomes on the test set. While the Standard LSTM demonstrates good performance, PmForecast consistently outperforms it, with distinctions becoming more pronounced for extended forecasting horizons beyond 12 hours. Following comprehensive experimentation and evaluation, the PmForecast model emerges as the most effective choice. PmForecast exhibits better performance achieving shallow RMSE values of 0.357 for 1-hour and 1.635 for 12-hour forecasts on the Carnot site test dataset. The temporal mechanism embedded in our PmForecast framework endows PmForecast with superior predictive capabilities among the evaluated algorithms, establishing it as a pivotal asset in our quest for precise PM predictions.

4.4 Multi-Tasks Model

Multi-task prediction offers efficiency, optimal resource utilization, and enhanced decision support, establishing itself as a valuable approach to air pollution prediction. Our research involved a comprehensive multi-task forecasting strategy, concurrently addressing the prediction of 3 major pollutants—PM_1, $PM_{2.5}$, and PM_{10}—as well as temperature and humidity. Figure 6 visually presents these correlations, illustrating the strong alignment between measured and predicted values for the next hour across the three PM fractions at the Canal station. While the single-task prediction model slightly outperformed the multi-task approach with a correlation of about 99% for $PM_{2.5}$ in one-hour prediction, it is important to highlight that the multi-task strategy using the PmForecast approach demonstrated remarkable efficiency in capturing the essence of the PM fraction's behavior with an evaluation metric for R^2 around 98% for all fractions. Equally impressive correlations were noted for other features, closely aligning with this value. Moreover, this robust performance was not limited to a specific dataset. Indeed, the correlation results remained consistent across the three additional sites further affirming the effectiveness of our multi-task forecasting methodology.

The overall efficiency and comprehensive insights provided by the multi-task approach, particularly with the PmForecast method, underscore its value and efficacy in capturing the complex behaviors of various PM fractions. This intriguing finding emphasizes the potential of the multi-task strategy as an effective alternative, providing comparable predictive accuracy while incorporating multiple parameters in the forecasting process.

4.5 Time Overhead for Model Training & Inferences

The PmForecast model displays variable time consumption across its key steps, as detailed in Table 4, utilizing an Apple M1 chip with 16GB of memory. The data pre-processing stage, involving one sample consisting of 5 measurements and 2 temporal features (day of the week and hour of the day), exhibits swift efficiency. Tasks include converting the timestamp from 15 seconds to hourly datasets and handling missing values using a moving average algorithm, achieving a latency of 61 seconds for the entire dataset in one station.

In contrast, the temporal mechanism, which operates on the entire dataset, introduces a favorable latency of less than 1 second, reflecting the simplicity associated with handling temporal aspects. The training phase for one epoch requires 5.0 milliseconds, emphasizing the computational demands involved in optimizing the model parameters. The complete model training process, from raw data to a trained model, takes 250 seconds, with a configuration of 200 epochs and an *early stopping* mechanism set for a patience of 30 epochs.

Recalibration configuration times vary depending on the horizon time. For instance, for 12-hour prediction points, the latency of a fully trained model is 283 seconds. Subsequently, predicting air pollution levels for a one-hour horizon demonstrates remarkable efficiency, with a latency of 375 milliseconds for 2400 samples. For the same number of samples but a longer horizon, such as 12 hours, the latency increases to 481 milliseconds.

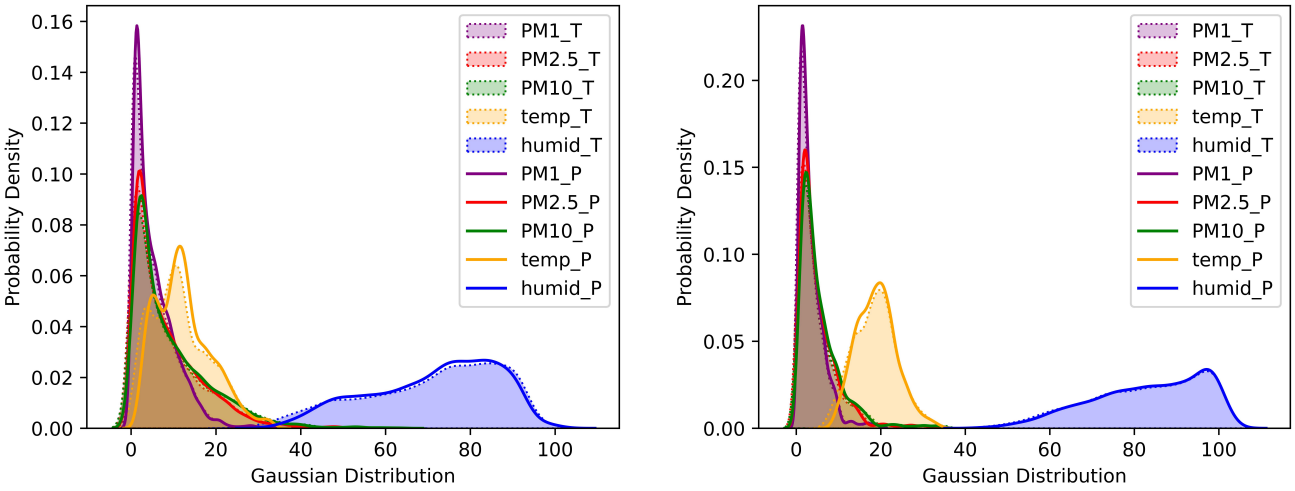
These temporal benchmarks offer insights into the computational performance of the PmForecast model, essential to assess its feasibility in online applications.

5 Conclusion

The study is dedicated to crafting an accurate air pollution predicting model, focusing primarily on $PM_{2.5}$. The wide spectrum of forecasting algorithms explored includes GPR, GRU, XGBoost, ARIMA, and Standard LSTM. In this comprehensive evaluation, the PmForecast model, an advanced iteration of LSTM,

Table 3: Assessing our Model's Predictive Performance at the Carnot Site Using Diverse Machine Learning Algorithms on the Test Dataset. **Bold values** indicate the best performance across all methods.

Forecasting Methods	1-Hour			12-Hours		
	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	R^2 (%)	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	R^2 (%)
GRU	0.424	0.240	98.2	1.648	1.022	70.0
GPR	0.615	0.351	96.2	3.182	1.495	11.0
XGBoost	0.437	0.237	98.1	1.874	1.083	65.1
ARIMA (VARMAX)	0.587	0.483	96.5	2.425	1.675	41.5
LSTM	0.422	0.256	98.2	1.731	1.019	70.2
PmForecast	**0.357**	**0.164**	**98.9**	**1.635**	**0.954**	**73.7**



(a) (b)

Fig. 6: Performance assessment through Gaussian distribution for multi-tasking at the Canal Site with varied meteorological data. (a) Examination of the correlation between observed and predicted values for the training set. (b) Investigation of the correlation between observed and predicted values for the test set. The truth and predicted values are illustrated with dotted and solid lines, featuring "T" and "P" in the labels, respectively. The colors represent the five measurements in our data: PM_1 (purple), $PM_{2.5}$ (red), PM_{10} (green), temperature (orange), and humidity (blue).

emerges as the unequivocal top performer, achieving superior accuracy rates of 98.9% and 73.7% for 1-hour and 12-hour forecasting periods, respectively. Through meticulous assessments conducted across various sites and forecasting horizons, PmForecast consistently proves to be the most effective choice, showcasing its unparalleled predictive capabilities.

In the context of multi-task predicting, which includes PM_1, $PM_{2.5}$, PM_{10}, temperature, and humidity, the single-task model slightly outperforms in $PM_{2.5}$ prediction. However, the PmForecast multi-task approach stands out for its remarkable efficiency, achieving high correlations of more than 98% for all PM fractions. This efficiency is particularly valuable for dynamic applications, as highlighted by varying time consumption metrics. PmForecast excels in terms of computational efficiency during data pre-processing and demonstrates low-latency predictions, reinforcing its potential for time-sensitive scenarios.

In conclusion, the PmForecast model not only stands out as a robust and versatile solution for accurate particulate matter prediction but also exhibits notable efficiency gains. Its implications for online monitoring and decision-making are underscored by its superior performance and computational efficiency, making it a valuable tool in air pollution forecasting applications.

Future Work

In the current and near future, our strategy involves deploying our lightweight air pollution forecasting model on a *Tensor Processing Unit* (TPU) microprocessor to achieve enhanced computational efficiency and enable *in situ* predictions by colocating the forecasting service and the sensors. We will implement regular *in situ* retraining strategies to keep the model updated with the latest data, ensuring continual relevance. Additionally, our focus is on developing an integrated spatio-temporal model that considers the interplay of diverse datasets, aiming for a more comprehensive understanding of air quality patterns.

Table 4: Time latencies for each step of the procedure in PmForecast.

Step	Latency (s)
Data pre-processing (8810 samples)	61
Temporal Mechanism (total data)	0.075
Training (1 epoch)	0.005
Train Full Model (full model from scratch to 1 point prediction for 1 station)	212
Online Model Recalibration (from 1 to 12 time points)	253
Prediction (1 time point for 2400 samples)	0.375
Inference Duration per Horizon (12 time points for 2400 samples)	0.481

Acknowledgement

This project has been financially supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie agreement No 847568. The research is conducted in collaboration with the *Centre de Recherche en Informatique, Signal et Automatique de Lille* (CRIStAL) in partnership with Inria of Lille and the *Laboratoire d'Optique Atmosphérique* (LOA), hosted at the University of Lille. The research is partly funded by INTPART (PACE partnership) which is a cooperation between The Research Council of Norway and research organizations in Norway, France, and Germany. Additionally, the quality-checked QameleO data for this project is sourced from the University of Bourgogne in the frame of the RESPONSE H2020 program. The funding for the QameleO network in Dijon came from the Ministry of Ecological Transition and Territorial Cohesion, and Dijon Metropolis in the frame of the POPSU PURE program. The installation of the QameleO network in Dijon was done in partnership between the University of Bourgogne and *Institut de Recherches pour le Développement* (IRD), co-owners of the QameleO instruments (open source license under Creative Common since 2020).

References

Abimannan S, Chang YS, Lin CY (2020) Air pollution forecasting using lstm-multivariate regression model. In: Hsu CH, Kallel S, Lan KC, et al (eds) Internet of Vehicles. Technologies and Services Toward Smart Cities. Springer International Publishing, Cham, pp 318–326, https://doi.org/https://doi.org/10.1007/978-3-030-38651-1_25

Alhirmizy S, Qader B (2019) Multivariate time series forecasting with lstm for madrid, spain pollution. In: 2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA), pp 1–5, https://doi.org/10.1109/ICCISTA.2019.8830667

Annesi-Maesano I, Forastiere F, Kunzli N, et al (2007) Particulate matter, science and eu policy. European Respiratory Society 29(3):428–431. https://doi.org/10.1183/09031936.00129506

Bui T, Le V, Cha S (2018) A deep learning approach for air pollution forecasting in south korea lstm. CoRR abs/1804.07891. https://doi.org/https://doi.org/10.48550/arXiv.1804.07891, URL http://arxiv.org/abs/1804.07891

Che Lah MS, Arbaiy N, Lin PC (2023) Arima-lp: A hybrid model for air pollution forecasting with uncertainty data. In: 2023 IEEE 8th International Conference On Software Engineering and Computer Systems (ICSECS), pp 353–356, https://doi.org/10.1109/ICSECS58457.2023.10256363

Cohen AJ, Brauer M, Burnett R, et al (2017) Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. Lancet (London, England) 389:1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6, URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5439030/

Crunaire S, Redon N, Spinelle L (2018) 1er essai national d'aptitude des micro-capteurs (eaɪc) pour la surveillance de la qualité de l'air: Synthèse des résultats. https://hal.archives-ouvertes.fr/hal-04250973, lCSQA. ffhal04250973f

Erbiao Y, Guangfei Y (2023) Sa–emd–lstm: A novel hybrid method for long-term prediction of classroom pm2.5 concentration. Expert Systems with Applications 230:120670. https://doi.org/https://doi.org/10.1016/j.eswa.2023.120670

Gangwar A, Singh S, Mishra R, et al (2023) The state-of-the-art in air pollution monitoring and forecasting systems using iot, big data and machine learning. ArXiv https://doi.org/https://doi.org/10.1007/s11277-023-10351-1

Guo Q, He Z, Li S, et al (2020) Air pollution forecasting using artificial and wavelet neural networks with meteorological conditions. Aerosol and Air Quality Research 20(6):1429–1439. https://doi.org/10.4209/aaqr.2020.03.0097, URL http://dx.doi.org/10.4209/aaqr.2020.03.0097

Guo Q, He Z, Wang Z (2023a) Predicting of daily pm2.5 concentration employing wavelet artificial neural networks based on meteorological elements in shanghai, china. Toxics 11(1). https://doi.org/10.

3390/toxics11010051, URL https://www.mdpi.com/2305-6304/11/1/51

Guo Q, He Z, Wang Z (2023b) Prediction of hourly pm2.5 and pm10 concentrations in chongqing city in china based on artificial neural network. Aerosol and Air Quality Research 23(6):220448. https://doi.org/10.4209/aaqr.220448, URL http://dx.doi.org/10.4209/aaqr.220448

Guo Q, He Z, Wang Z (2023c) Simulating daily pm2.5 concentrations using wavelet analysis and artificial neural network with remote sensing and surface observation data. Chemosphere 340:139886. https://doi.org/https://doi.org/10.1016/j.chemosphere.2023.139886, URL https://www.sciencedirect.com/science/article/pii/S0045653523021550

Guo Q, He Z, Wang Z (2024) The characteristics of air quality changes in hohhot city in china and their relationship with meteorological and socio-economic factors. Aerosol and Air Quality Research 24(5):230274. https://doi.org/10.4209/aaqr.230274, URL http://dx.doi.org/10.4209/aaqr.230274

Harrison P (2016) The air we breathe. https://doi.org/10.1183/09031936.93.01050397, URL https://www.researchgate.net/publication/295626722_The_air_we_breathe

He J, Li X, Chen Z, et al (2023) A hybrid clstm-gpr model for forecasting particulate matter (pm2.5). Atmospheric Pollution Research 14(8):101832. https://doi.org/https://doi.org/10.1016/j.apr.2023.101832

He Z, Guo Q, Wang Z, et al (2022) Prediction of monthly pm2.5 concentration in liaocheng in china employing artificial neural network. Atmosphere 13(8). https://doi.org/10.3390/atmos13081221, URL https://www.mdpi.com/2073-4433/13/8/1221

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural computation 9(8):1735–1780. https://doi.org/https://doi.org/10.1162/neco.1997.9.8.1735

Hu W, Fang L, Zhang H, et al (2023) Changing trends in the air pollution–related disease burden from 1990 to 2019 and its predicted level in 25 years. Environmental Science and Pollution Research 30:1761–1773. https://doi.org/10.1007/s11356-022-22318-z, URL https://link.springer.com/article/10.1007/s11356-022-22318-z

Le P, Zuidema WH (2016) Quantifying the vanishing gradient and long distance dependency problem in recursive neural networks and recursive lstms. In: Proceedings of the 1st Workshop on Representation Learning for NLP, Rep4NLP@ACL 2016, Berlin, Germany, August 11, 2016. Association for Computational Linguistics, pp 87–93, https://doi.org/10.18653/V1/W16-1610, URL https://doi.org/10.18653/v1/W16-1610

Lee HM, Park RJ, Henze DK, et al (2017) Pm2.5 source attribution for seoul in may from 2009 to 2013 using geos-chem and its adjoint model. Environmental Pollution 221:377–384. https://doi.org/https://doi.org/10.1016/j.envpol.2016.11.088

Li J, An X, Li Q, et al (2022) Application of xgboost algorithm in the optimization of pollutant concentration. Atmospheric Research 276:106238. https://doi.org/https://doi.org/10.1016/j.atmosres.2022.106238

Manisalidis I, Stavropoulou E, Stavropoulos A, et al (2020) Environmental and health impacts of air pollution: A review." frontiers in public health. Front Public Health 8(14). https://doi.org/10.3389/fpubh.2020.00014

Martiny N, Nicolas M, Sarah M, et al (2023) Quality of air module for environmental learning engineering and observation network (qameleondijon) : un réseau dense de mesures de qualité de l'air à dijon», climatologie 20(4). https://doi.org/https://doi.org/10.1051/climat/202320004

Mikati I, Benson AF, Luben TJ, et al (2018) Disparities in distribution of particulate matter emission sources by race and poverty status. American Journal of Public Health 108:480–485. https://doi.org/10.2105/AJPH.2017.304297, URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5844406/

Ngom B, Diallo M, Seyc MR, et al (2021) Pm10 data assimilation on real-time agent-based simulation using machine learning models: case of dakar urban air pollution study. In: 2021 IEEE/ACM 25th International Symposium on Distributed Simulation and Real Time Applications (DS-RT), pp 1–4, https://doi.org/10.1109/DS-RT52167.2021.9576143

Noh S (2021) Analysis of gradient vanishing of rnns and performance comparison. Inf 12(11):442. https://doi.org/10.3390/INFO12110442, URL https://doi.org/10.3390/info12110442

Panneerselvam V, Thiagarajan R (2023) Acbigru-dao: Attention convolutional bidirectional gated recurrent unit-based dynamic arithmetic optimization for air quality prediction. Environmental Science and Pollution Research 30(37):86804–86820. https://doi.org/10.1007/s11356-023-28028-4, URL https://doi.org/10.1007/s11356-023-28028-4

Philip JL, Richard F, Nereus JRA, et al (2017) The lancet commission on pollution and health. Lancet https://doi.org/10.1016/S0140-6736(17)32345-0, URL http://dx.doi.org/10.1016/S0140-6736(17)32345-0

Ponomarev N, Elansky N, Kirsanov A, et al (2020) Application of atmospheric chemical

transport models to validation of pollutant emissions in moscow. Atmospheric and Oceanic Optics 33:362–371. https://doi.org/https://doi.org/10.1134/S1024856020040090

Powers JG, Klemp JB, Skamarock WC, et al (2017) The weather research and forecasting model: Overview, system efforts, and future directions. Bulletin of the American Meteorological Society 98(8):1717–1737. https://doi.org/https://doi.org/10.1175/BAMS-D-15-00308.1

prabhu V, Shridhar V (2019) Investigation of potential sources, transport pathway, and health risks associated with respirable suspended particulate matter in dehradun city, situated in the foothills of the himalayas. Atmospheric Pollution Research 10(1):187–196. https://doi.org/https://doi.org/10.1016/j.apr.2018.07.009, URL https://www.sciencedirect.com/science/article/pii/S1309104218302770

Redon N, Crunaire S, Herbin B, et al (2018) French joint intercomparison exercises for air quality sensors (eaµc): Results and assessment. In: International Symposium on Individual air pollution sensors: Innovation or Revolution

Srivastava H, Kumar Das S (2023) Air pollution prediction system using xrsth-lstm algorithm. Environmental Science and Pollution Research 30(60):125313–125327. https://doi.org/10.1007/s11356-023-28393-0

Terzano C, Stefano F, Conti V, et al (2010) Air pollution ultrafine particles: Toxicity beyond the lung. European review for medical and pharmacological sciences 14:809–21. https://doi.org/https://doi.org/10.1016/j.gsf.2021.101147

Tsai YT, Zeng YR, Chang YS (2018) Air pollution forecasting using rnn with lstm. In: 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), pp 1074–1079, https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00178

Tu Z, Wu Z (2022) Predicting beijing air quality using bayesian optimized cnn-rnn hybrid model. In: 2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML), pp 581–587, https://doi.org/10.1109/CACML55074.2022.00104

Verma I, Ahuja R, Meisheri H, et al (2018) Air pollutant severity prediction using bi-directional lstm network. pp 651–654, https://doi.org/10.1109/WI.2018.00-19

Wang J, Bai L, Wang S, et al (2019) Research and application of the hybrid forecasting model based on secondary denoising and multi-objective optimization for air pollution early warning system. Journal of Cleaner Production 234:54–70. https://doi.org/https://doi.org/10.1016/j.jclepro.2019.06.201

World Health Organization (2006) Air quality guidelines: global update 2005. https://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/, accessed: October 16, 2023

Zaree T, Honarvar A (2018) Improvement of air pollution prediction in a smart city and its correlation with weather conditions using metrological big data. Turkish Journal of Electrical Engineering and Computer Sciences 26(3):1302 – 1313. https://doi.org/10.3906/elk-1707-99

Zhang B, Wu B, Liu J (2020) Pm2.5 pollution-related health effects and willingness to pay for improved air quality: Evidence from china's prefecture-level cities. Journal of Clean 273:122876. https://doi.org/10.1016/j.jclepro.2020.122876, URL https://www.sciencedirect.com/science/article/pii/S0959652620329218

Zhang Q, Han Y, Li VOK, et al (2022a) Deepair: A hybrid cnn-lstm framework for fine-grained air pollution estimation and forecast in metropolitan cities. IEEE Access 10:55818–55841. https://doi.org/10.1109/ACCESS.2022.3174853

Zhang Z, Zhang S, Zhao X, et al (2022b) Temporal difference-based graph transformer networks for air quality pm2.5 prediction: A case study in china. Frontiers in Environmental Science 10. https://doi.org/https://doi.org/10.3389/fenvs.2022.924986

Zhao J, Luo H, Sang W, et al (2023) Spatiotemporal semantic network for ENSO forecasting over long time horizon. Appl Intell 53(6):6464–6480. https://doi.org/10.1007/S10489-022-03861-1, URL https://doi.org/10.1007/s10489-022-03861-1

Zhu H, Hu J (2019) Air quality forecasting using svr with quasi-linear kernel. In: 2019 International Conference on Computer, Information and Telecommunication Systems (CITS), pp 1–5, https://doi.org/10.1109/CITS.2019.8862114

Declarations