Development of Models for Air Pollution-Related Public Health Assessment: Application of Long Short-Term Memory Neural Network for Short-term Exposure Effect

A Thesis Submitted to the Committee on Graduate Studies

in Partial Fulfillment of the Requirements for the Degree of Master of

Science in the Faculty of Arts and Science

TRENT UNIVERSITY

Peterborough, Ontario, Canada

Applied Modeling and Quantitative Methods M.Sc. Graduate Program

September 2025

Abstract

Development of Models for Air Pollution-Related Public Health Assessment: Application of Long Short-Term Memory Neural Network for Short-term Exposure Effect

Huawei Han

This thesis develops an Long Short-Term Memory (LSTM) neural network model to assess the relationship between ambient air pollutant exposure and public health risks, accommodating both linear and nonlinear associations with distributed lags. The research makes three key contributions. **First**, Maximal Information Coefficient (MIC) methods identify the most relevant air pollutants and their associations with health outcomes. **Second**, an LSTM model extracts temporally dependent features from exposure series to estimate health impacts. **Finally**, the model's potential in air pollution epidemiology is explored using Local Interpretable Model-Agnostic Explanations (LIME) to interpret the exposure-health response relationship.

Keywords: air pollution epidemiology, public health assessment, Maximal Information Coefficient, Long Short-Term Memory, Local Interpretable Model-Agnostic Explanations

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Wesley Burr for his enthusiasm, encouragement and guidance throughout this research work. His insight and expertise in statistics inspired me a lot in the completion of this thesis. And my thanks to Dr. Hwashin Shin, my co-supervisor on this project, for providing me access to Health Canada's data and models. I also thank David Riegert for his encouragement and insightful comments.

Second, my heartfelt thanks also go to George and Marilyn MacDonald, Mark and Ellen Stephen, Norm and Priscilla Tucker, Frank and Jane Hamilton, Carolina, Beverly, and so many other friends from the Riverside Community for their warmth, kindness, inclusiveness, and support during my study at Peterborough.

Lastly, I thank my family for their unconditional love and encouragement. This work would not have been possible without their patience and support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CO carbon monoxide

NO nitric oxide

NO_2 nitrogen oxide

NO_x nitrogen oxides

O_3 ground ozone

PM_{10} particulate matter less than or equal to 10 micrometers

$PM_{2.5}$ particulate matter less than or equal to 2.5 micrometers

SO_2 sulphur dioxide

T daily mean temperature

ADAM Adaptive Moment Estimation

APHEA Air Pollution and Health: a European Approach

GAM Generalized Additive Model

GLM Generalized Linear Model

LIME Local Interpretable Model-Agnostic Explanations

LSTM Long Short-Term Memory

MAE Mean Absolute Error

MAPE Mean Absolute Percentage Error

MI Mutual Information

MIC Maximal Information Coefficient

NAPS National Air Pollution Surveillance

NMMAPS National Morbidity and Mortality Air Pollution Study

PCC Pearson Correlation Coefficient

R^2 Explained Variation

RMSE Root Mean Square Error

RNN Recurrent Neural Network

Chapter 1

Introduction

1.1 Background

Short- and long-term exposure to air pollutants has been associated with various adverse public health effects in numerous epidemiological studies. Exposure to major regulated air pollutants like ground-level ozone, particulate matters, oxides of sulphur and oxides of nitrogen has become recognized as one of the high environmental risks and has been shown to have a relationship with increasing cardiovascular disease [1, 2, 3, 4], respiratory disease [5, 6, 7], and public mortality and morbidity [8, 9, 10]. In order to monitor and improve air quality, learn about the health impacts of air pollutants, and reduce the burden of social healthcare, a series of projects have been launched in several countries or regions to research ambient air quality monitoring and the statistical relationship between air pollution exposure and specific health outcomes.

Among these air pollution-related air quality and public health studies, three

influential programs are worth mentioning. The first one is the National Morbidity and Mortality Air Pollution Study (NMMAPS) [11, 12], which was initiated in 1996 with the main purpose of studying the association between particulate matters and daily mortality, as well as the adverse effects of several regulated air pollutants on public health in United States. This research work was conducted by investigators from Johns Hopkins University and Harvard University, and was funded and supported by the Health Effects Institute. Air pollution data and daily mortality of 108 large cities from 1987 to 2000 in the United States were collected. Time-series analysis of several air pollutants and association between air pollutants and daily mortality were also explored. The second is the Air Pollution and Health: a European Approach (APHEA) project, including APHEA and APHEA2 [13, 14, 15]. These projects began in 1993, which also focused on investigating the adverse effects of short-term exposure to particulate matter and other air pollutants on public health in European cities. Quantitative evaluations of their short-term effects using data from 15 cities in APHEA and 8 in APHEA2 were investigated, where associations between exposure to ambient particles and mortality or hospital admissions were identified. The third such project is the National Air Pollution Surveillance (NAPS) Program [16, 17] started in 1969, which is managed by Environment and Climate Change Canada and has multiple purposes. One of these, co-sponsored by Health Canada, includes monitoring the variation of the Air Quality Health Index and Canadian Environmental Sustainability Indicators. Ambient air pollution data of NAPS are collected from more than 250 monitoring stations distributed across

Canada and are publicly accessible through the Canada-Wide Air Quality Database.

The above-mentioned programs have been promoting the development of quantitative and statistical methods for adverse impacts assessment of air pollution on public health, and achieved significant results [18, 19]. Collected historical data of air pollution and medical records and implementation of innovative statistical tools in these research are still having a far-reaching influence on current air pollution-related epidemiology studies.

1.2 Literature Review

One of the major topics in air pollution epidemiology research is to investigate the impacts of short- or long-term exposure to air pollutants of interest on specific public health outcome. Various statistical methods and epidemiological designs have been explored for this problem since the mid-twentieth century. These research methods generally fall into two categories: cross-sectional studies and longitudinal studies, with the former investigating air pollutants exposure and health outcome at a single point in time and the latter over a period of time [20, 21].

In earlier studies, as systematic monitoring and data collection were not available, miscellaneous methods under the above-mentioned two categories were explored. Most of the research were conducted for air pollution-caused health incidents or based on observation of specifically designed experiments, where population groups, types of pollutants, exposures, etc., were regulated. In 1961, Ciocco *et al.* [22] investigated the impacts of air pollution on residents in Donora and Pittsburgh

using data from community surveys ten years after the Donora smog disaster in 1948 and showed higher morbidity and mortality were related to exposure to higher air pollutant concentration. In 1970, multiple regression analysis method was introduced to quantitatively estimate the long-term effects of air pollutants using data from particular locations of England and showed a strong relationship between specific air pollutants and several diseases [23]. Similar methods were also used later by Lipfert [24], Gibbons and McDonald [25] and Ostro [26].

Later, with the need to investigate specific health outcome caused by exposure to air pollutants of interest over a period of time, cohort studies were used with prospective cohorts investigating future health outcomes and retrospective cohorts for past relationship. In 1973, a prospective cohort study was conducted by Waller *et al.* [27] to investigate the lasting impacts of exposure to London fog in the 1950s, and higher prevalence of respiratory symptoms in the exposed residents was reported. From 1978 to 1981, Kerigan *et al.* [28] conducted a three-year cohort study to investigate the influence of ambient air pollution on children's respiratory health in Hamilton, Ontario. This method was also applied to a series of air pollution epidemiology studies such as cancer [29, 30, 31], diabetes [32, 33], cardiovascular disease [34, 35], mortality [36, 37, 38], etc. Panel studies are another longitudinal research design similar to cohort studies used in air pollution epidemiology, with the main difference that the former focuses on the same group of population while the latter on population with shared characteristics. Air pollution panel studies were also used in a variety of epidemiological problems, such as asthma [39, 40], cardiac

disease [41], pulmonary disease [42, 43], neurobehavioral disorder [44], inflammation [45, 46], and so on.

In addition, case-crossover analysis is also an extensively used longitudinal method in air pollution epidemiology when estimating their acute impacts on same or similar population group through contrasting before and after some health outcomes [47]. Since the development of the case-crossover method in 1991 [48], it has been applied to a variety of experimental designs for air pollutants-related epidemiological studies. In 1999, case-crossover designs were made for assessing the effects of air pollution on mortality in Philadelphia, Pennsylvania, USA [49] and Seoul, Korea [50]. The effects of acute exposure to air pollution on sudden cardiac arrest and asthma attack were investigated by Checkoway *et al.* [51] and Im *et al.* [52] in 2000. In the following decades, case-crossover analysis was used in impacts assessment of air pollution on myocardial infarction [53] (2003), coronary mortality [54] (2005), otitis media [55] (2010), respiratory disease [56] (2014) and incident pneumonia [57] (2018).

With development of air pollutant monitoring systems and availability of continuously collected data, longitudinal analysis using time series methods were applied to investigate the cumulative impacts of air pollution exposure. Two extensively discussed models are Generalized Linear Model (GLM)s and its extension Generalized Additive Model (GAM)s, where the health outcome is modeled using Poisson response variables as a function of air pollution exposure. In 1979, a logistic regression model was presented by Korn and Whittemore [39] to analyze the impacts of exposure to varying concentrations of air pollution on asthma, with the motivation

to overcome the problems of independence requirement and missing response data in linear regression. In the following years of time series study for air pollution epidemiology, Poisson regression was proposed to analyze a series of air pollutants-related epidemiological issues like morbidity [58], respiratory disease [59, 60], lung cancer [61] and mortality [62, 63].

Around 2000, GLMs and GAMs were extensively applied to air pollution epidemiology by a number of researchers, where quite a few significant works were motivated by the aforementioned three programs. In 2000, Dominici *et al.* [18] presented a general semiparametric log-linear regression model to estimate the PM_{10}-associated mortality rate in 20 large US cities and a hierarchical regression model was also developed to investigate the spatial correlation of responding coefficients among different cities. In the same year, Sheppard and Damian [64] proposed a Poisson regression model for health effects assessment of short-term air pollutant exposure, which combined exposure distribution and measurement error with the disease model. This model was applied to particulate matter-associated asthma hospital admission in Seattle and showed that the measurement error did not affect the results of the time series regression model. Similar methods were also used in APHEA projects. In 1998, Spix *et al.* [65] used Poisson regression models and standard confounder models to investigate the air pollution-related hospital admissions of different age groups in five West Europe cities and showed that ozone had a significant impact on respiratory diseases. In 2001, researchers from APHEA2 project [66] applied Poisson regression to particle matter-related mortality analysis in 29 European cities, taking

advantage of its non-parametric smoothing of seasonal patterns, and confirmed the impacts of ambient particles on mortality while finding the characteristics-related heterogeneity of effect parameters.

Besides these single-pollutant time series models, multi-pollutant models under the GAM framework were also investigated. In 2004, Zeka and Schwartz [67] presented a two-pollutant Poisson regression model by linearly correlating two kinds of air pollutants to investigate the effects of particulate matters and several other air pollutants on daily mortality based on the NMMAPS datasets. They confirmed the effects of PM_{10} on daily mortality and reported a carbon monoxide-related daily mortality increase. In order to estimate the cumulative effects of exposure, Zanobetti *et al.* [68] developed GAMs with distributed lags that related some health outcome on a given day with air pollution exposure in prior days. Timescale effects of exposure were also investigated by Dominici *et al.* [69] by using predictors from Fourier decomposition of air pollution time series in GAM and later by Burr *et al.* [70] through combining the lagged predictors in GAM. Currently, GAMs are still one of the most popular time series methods used in air pollution epidemiology, with application in multi-site time series [71], bias correction [72], multi-pollutant effects [73, 74], temporal trends [75, 76] and many more applications.

In summary, longitudinal study has become the mainstream method in air pollution epidemiology at present, of which time series methods, especially GAMs, are widely used. Moreover, complex epidemiological designs with combination of time series and other longitudinal methods have been gradually developed for meticulous

investigation of the health effects from cumulative air pollution exposure.

1.3 Challenges of the Problem

As the most widely used time series model in air pollution epidemiology, GAMs provide a perspicuous air pollution exposure-health response assessment framework taking advantage of Poisson regression for various health events, non-parametric splines for exposure effects estimation and confounding elements smoothing, while the following issues and challenges still exist in its epidemiological applications.

1. Potentially biased estimates for multi-pollutant GAMs: different air pollutants are correlated with each other. Air pollutants are usually mixed with each other physically and chemically from their generating sources to spreading, especially for some strongly related ones like different oxides of nitrogen or particulate matters with different diameters. As people are generally exposed to multiple air pollutants simultaneously, finding the primary pollutant for some particular heath outcome is usually not easy in the GAM-based modelling and analysing process. Capturing and evaluating the mixed health impacts of correlated air pollutants by smoothing the confounding effects may not be enough and keep an open question for further investigation[77]. In addition, collinearity (also known as multicollinearity) brought by the associations reduces the interpretability of response coefficients for air pollutant of interest and weakens the power of GAMs. We try to address this issue by introducing MIC to measure both the linear and non-linear associations in chapter 2, as well as by proposing LSTM models in chapter 3 with multiple air

pollutants included and their non-linear associations considered.

2. Reduced sensitivity of health outcomes to air pollution: performance of the GAMs is limited. GAMs present a straightforward exposure-response relationship by fitting the model using time-varying air pollution concentrations and health events data, which facilitates analysis of the degree to which health outcomes respond to air pollution variations. However, in most air pollution epidemiological applications, to what extent the fitted model can reflect the true relationship between air pollutants of interest and particular health outcome is generally not evaluated. Either underfitting or overfitting may degrade its performance of generalization, and thus reliability of the exposure-response coefficients obtained from the GAMs might need further examination. In chapter 3, with the proposed LSTM models we analyzed the fitting performance of both the new models and traditional GAMs.

3. Limited integration of lagged exposure-response relationships: public health outcomes are affected by cumulative effects of exposure. There are plenty of air pollution epidemiology studies assumed that an adverse health outcome at one time period is associated with exposure at a previous period. This kind of formulation is referred to as single lag model, while its alternative is distributed lag model, which involves cumulative effects of exposures at multiple time periods before the health event occurs and conforms better to more realistic situations. Parametric designs are usually used to model the relationship between previous distributed exposures and lagged responses. Most of these formulations focus on the cumulative effects of a single air pollutant for parameters fitting and coefficients interpretation issues[70].

Further studies are still needed to investigate the effects of exposure to multiple air pollutants in distributed lag models. In the method we proposed with LSTM models, we apply the LIME approach in chapter 5 for analyzing the lag effects on the health outcomes.

1.4 Research Framework

Aiming at the above-mentioned issues, the following main research work related to quantitative health risk analysis in air pollution epidemiology is made in this thesis.

1) For issue 1, Mutual Information (MI)-based Criterions (MIC) are used to evaluate the association between different air pollutants and find the most related factors for health outcome of interest, taking advantage of information entropy to capture both the linear and nonlinear relation.
2) For issues 1 and 2, an LSTM model is developed to assess the impacts of exposure to multiple air pollutants on health outcomes of interest with weighted evaluation of distributed lags. An LSTM neural network is first designed to extract health outcome-related feature information from multipollutant exposure sequence with temporal dependence. Then estimation layers with weighted evaluation of the extracted features from the exposure that has distributed lags are constructed to assess the health outcome of interest.
3) For issue 3, interpretability of the LSTM model using LIME is investigated to analyze feature importance, examining predictors' temporal significance and the relationship between PM10 and temperature in mortality outcome estimation.

The other parts of this thesis are organized as follows. In Chapter 2, MIC-based association analysis and air pollutants selection for health outcome of interest are presented. In Chapter 3, the LSTM model with weighted evaluation of the air pollution-related health outcome is developed. In Chapter 4, preliminary exploration of the exposure-response relation based on the proposed LSTM model is performed and potential of the LSTM model in air pollution-related public health assessment using LIME is demonstrated. At last, a brief conclusion is made and potential future research opportunities are discussed in Chapter 5.

Chapter 2

Correlation of Air Pollutants

In this chapter, datasets used for analysis throughout this thesis are first introduced. Then Mututal Information (MI) and Maximal Information Coefficient (MIC) are applied to correlation analysis between different air pollutants for air pollution variables selection. Part of this chapter is published in the 6th International Conference on Statistics: Theory and Applications (ICSTA 2024) [78].

2.1 Datasets Description

The ambient air pollutants data used throughout this work are mainly from the publicly available Canada-Wide Air Quality Database with pollution data from the NAPS program and the database for NMMAPS. The background of these two databases are briefly introduced in the following.

The NAPS database includes continuous and hourly measurements of nitric oxide (NO), nitrogen oxide (NO_2), nitrogen oxides (NO_x), carbon monoxide (CO),

sulphur dioxide (SO_2), ground ozone (O_3), particulate matter less than or equal to 2.5 micrometers ($PM_{2.5}$) and particulate matter less than or equal to 10 micrometers (PM_{10}). These data are available from the Government of Canada Open Data Portal [79] and detailed description of the program and datasets is available at Open Government [80]. As the data availability is reduced due to the impact of the COVID-19 pandemic starting in 2021, 22 years of air pollution data from 2000 to 2021 are collected. In addition, daily mean temperature (T) for the same periods are collected from the Canadian Centre for Climate Services [81]. All categories of data are preprocessed as follows to facilitate the numerical experimentation. The raw air pollution time series data in the NAPS datasets are measured on an hourly basis and their 24-hour means are first calculated since all the analyses throughout this work are conducted on a daily basis. Then a second-order polynomial interpolation is performed for the missing values. At last, the outliers including negative values of the interpolated time series are replaced using the mean of values before and after the interpolated day.

Daily mean temperature and seven air pollutants data from 2000-2021 for Toronto, ON, Canada are shown as an example in Figure 2.1. It can be roughly seen that T and O_3 have obvious yearly periodicity. People are prone to be exposed to higher O_3 concentrations in higher temperature periods, as O_3 is closely related to solar irradiance in its formation process. Another obvious point is that daily concentrations of CO, NO, NO_2, NO_x and SO_2 all have downward trends in the given periods, possibly due to various stricter air pollution emission standards. It can also be seen

that the $PM_{2.5}$ series stays relatively stable over this time span. The PM_{10} data are excluded for the reason that the vast majority of them are missing.

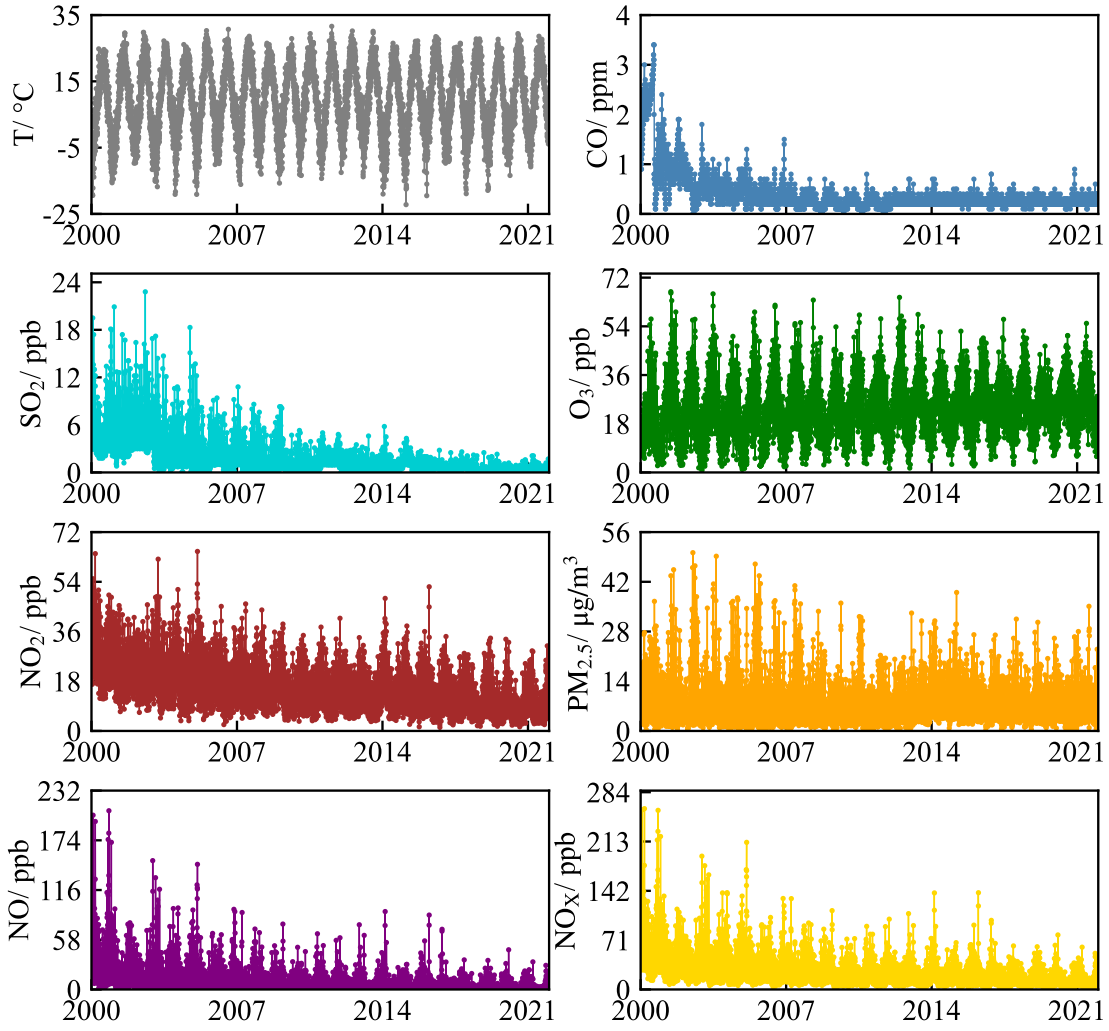

Figure 2.1: Daily mean values of temperature and air pollutant concentration for Toronto, ON, Canada from 2000 to 2021.

The NMMAPS database includes daily measurements of NO_2, CO, SO_2, O_3, $PM_{2.5}$ and PM_{10}. The air pollution data, weather condition and health outcomes are assembled and organized by city in an R package `NMMAPSdata` [82, 83]. As the updated data after 2000 were never made publicly accessible, the time periods used for analysis throughout this work are from 1987 to 2000. Data from Chicago

was selected for our investigation, which includes the daily mean temperature (T), CO, SO_2, O_3, NO_2, PM_{10} and daily non-accidental mortality from several diseases. The data used for investigation are visualized and shown in Figure 2.2 after being processed as follows:

1) as the original temperature data from Jan. 1, 1998 to Dec. 31, 2000 are missing, online weather data of Chicago [84] for the three years are collected and supplemented (using the Fahrenheit scale).
2) the original air pollutants data series are detrended by subtracting a 365 day moving average. They are restored to the true monitor values by adding the trimmed mean and the corresponding trend item [82] to facilitate our following analysis.
3) daily non-accidental mortality data for three age groups (under 65, 65 to 74, and above 75) are aggregated as the daily morality observation.
4) other small portions of missing values are also interpolated and outliers are dealt with as the above NAPS datasets.

It can also be seen in Figure 2.2 that T and O_3 data series have obvious yearly periodicity as in the NAPS datasets. The daily all-cause non-accidental mortality data series show roughly regular fluctuations on a yearly basis as well. Compared to the NAPS datasets, other air pollutants series present no evident periodicity or downward trends. In addition, $PM_{2.5}$ data are not included as most of them are missing.

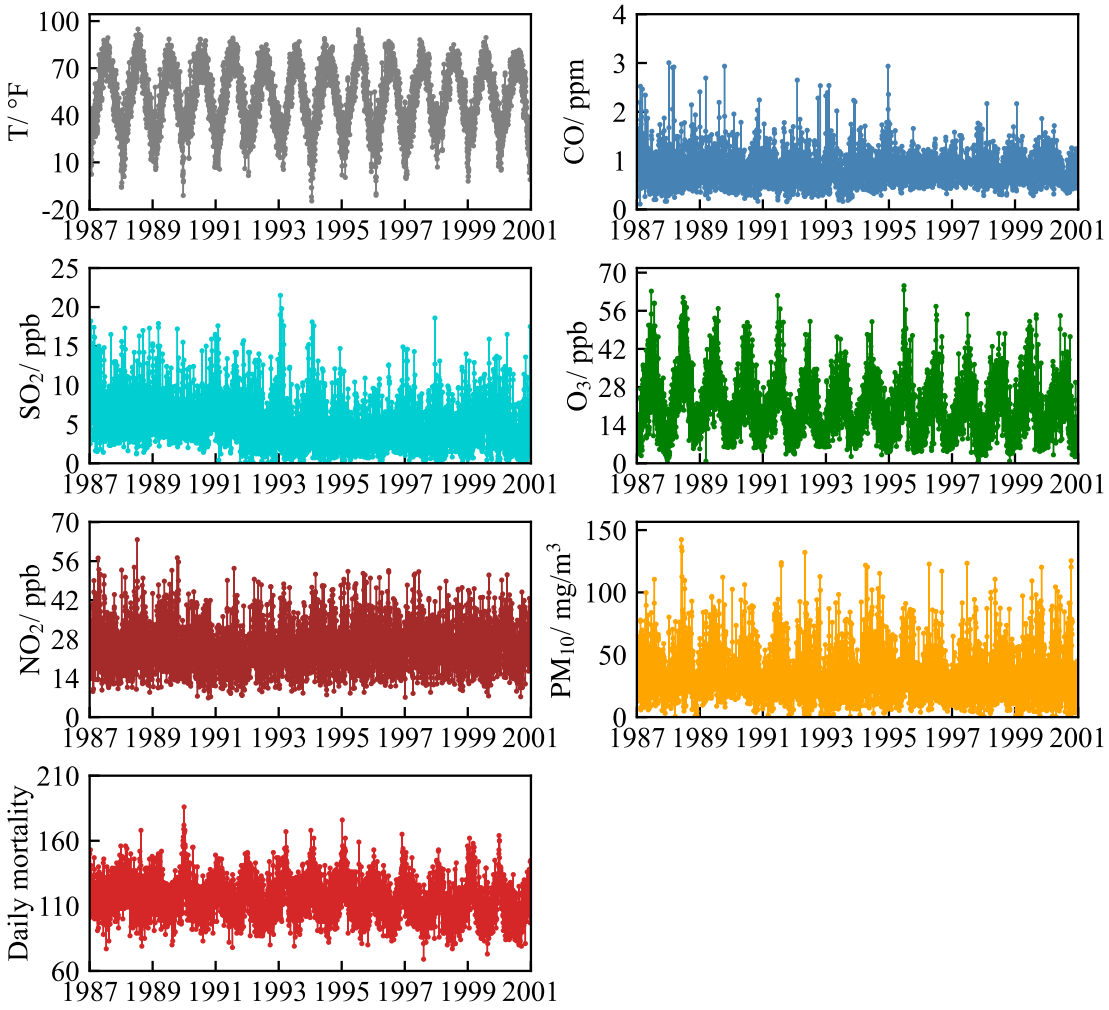

Figure 2.2: Daily mean values of temperature and air pollutants concentration and daily mortality for Chicago, IL, the U.S. from 1987 to 2000.

2.2 Mutual Information Based Association

In air pollution epidemiology studies, as different air pollutants are physically and chemically mixed with each other in diffusion process and people are generally exposed to different air pollutants at the same time, health impacts assessment models with multiple air pollutants are more realistic. Thus, correlation analysis for different air pollutants and weather conditions is usually done to identify the most significant predictor(s) and facilitate the modelling process. However, associations between dif-

ferent predictors are not always linear in various health assessment problems, and hence the Pearson Correlation Coefficient (PCC) used for relationship analysis and predictors screening [85] is not always appropriate. The relationship between several air pollutants of Toronto is shown as an example in Figure 2.3. The relation between NO and NO_x is approximately linear, while the linear relation between O_3 and temperature, O_3 and $PM_{2.5}$ as well as NO_2 and CO is not obvious. Using PCC to evaluate their relevance may not be able to capture their true association, especially when choosing the air pollutants of interest that contribute to specific health problems.

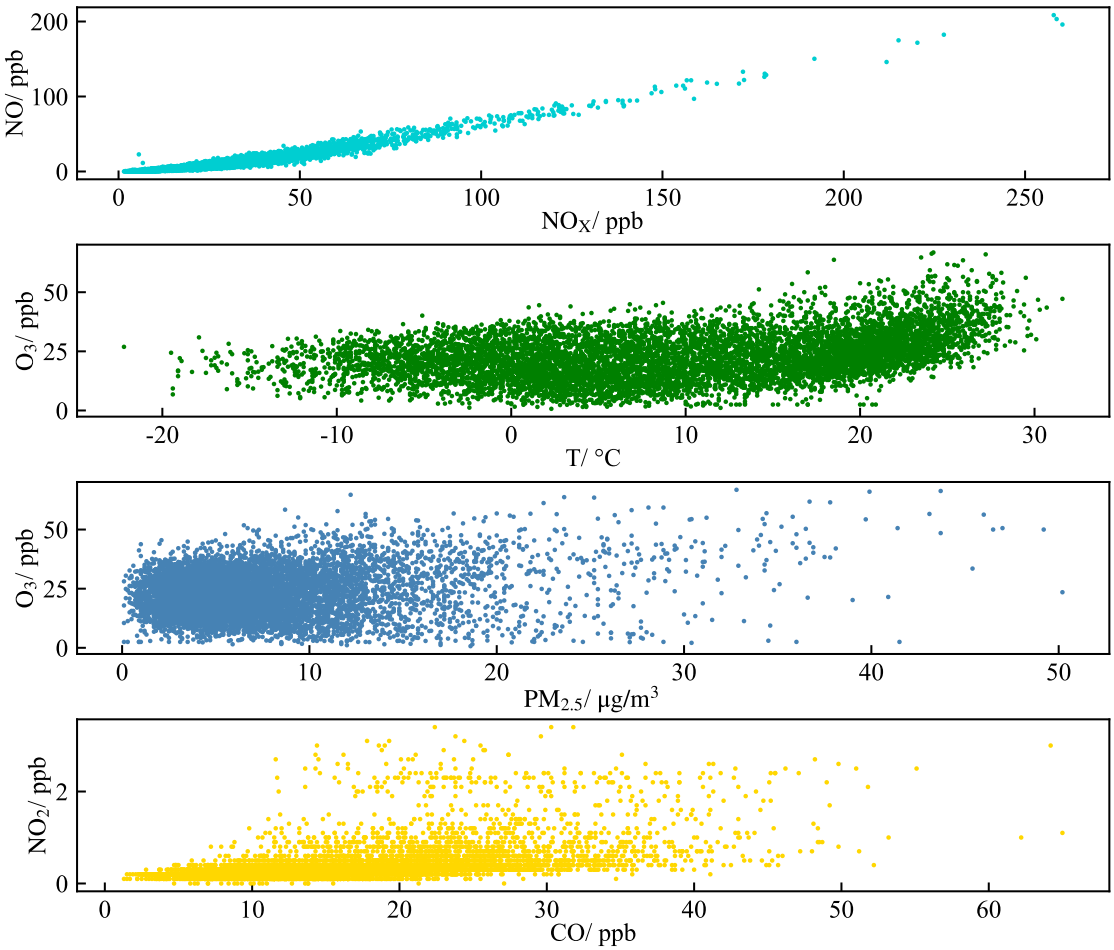

Figure 2.3: Relations of different air pollutants for Toronto, ON, Canada.

In order to evaluate the strength of association between different health predictors in a general form, Mututal Information (MI) is used to quantify both the linear and nonlinear relation in this work. MI is a measure of strength in statistics that evaluates the mutual dependence of two random variables, which quantifies how much information for one random variable can be obtained after observing another, using entropy [86]. For two different air pollution variables X_1 and X_2, their MI is expressed as

$$MI(X_1, X_2) = H(X_1) - H(X_1|X_2), \tag{2.1}$$

where $H(X_1)$ is the information entropy for X_1. $H(X_1|X_2)$ is the conditional information entropy of X_1 after observing X_2, which can be expressed as:

$$H(X_1|X_2) = H(X_1, X_2) - H(X_2), \tag{2.2}$$

where $H(X_2)$ is the information entropy for X_2 and $H(X_1, X_2)$ is the joint entropy of X_1 and X_2. The information entropy formula [87] for a random variable X with discrete distribution is

$$H(X) = \sum_x p_X(x) \log_2 p_X(x)^{-1}, \tag{2.3}$$

where $p_X(x)$ is the probability of taking value x for variable X.

With Eqn. (2.2), Eqn. (2.3) and conditional probability formula, MI in Eqn. (2.1) can be further formulated using probabilities as

$$MI(X_1, X_2) = \sum_{x_1} \sum_{x_2} p_{X_1X_2}(x_1, x_2) \log_2 \left[\frac{p_{X_1X_2}(x_1, x_2)}{p_{X_1}(x_1) p_{X_2}(x_2)} \right], \tag{2.4}$$

where $p_{X_1X_2}(x_1, x_2)$ is the joint probability distribution of X_1 and X_2, and $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$ are their marginal distributions. A larger MI value indicates a stronger

association between X_1 and X_2, while a lower one indicates that they are weakly associated. Zero value means that they have no statistical association.

2.3 Maximal Information Coefficient for Air Pollutants Selection

MI in Eqn. (2.4) can capture both linear and nonlinear associations between different air pollutants, although there still exist some limitations when applied to our air pollution datasets. The first issue is that the air pollutants data are measured using different metrics with units on multiple scales. Therefore, MIs between different variables may have various scales, with which comparison of MIs is not possible. The second is that calculation of the joint probability distribution in Eqn. (2.1) is time-consuming and difficult. Therefore, Maximal Information Coefficient (MIC) is used for association evaluation instead, which is a quantity developed from MI to measure the association between two random variables in a normalized form. Its main idea is to use many possible binning methods to partition the scatter plot of variables into different bins and calculate the MIs with estimation of the joint probability distribution accordingly. Then the gridding method with the maximal normalized MI is chosen as the MIC formulation shown in Eqn. (2.5),

$$MIC(X_1, X_2) = \max_{n_{X_1} \times n_{X_2} \leq N} \frac{MI(X_1, X_2)}{\log_2 \min(n_{X_1}, n_{X_2})}, \tag{2.5}$$

where n_{X_1} and n_{X_2} are the numbers of partitions for pairwise air pollutants data of X_1 and X_2, and N is a parameter related to the sample size S and is set to $S^{0.6}$ in

the following experiments. The range of $MIC(X_1, X_2)$ falls into $[0, 1]$, with which the association between pairwise air pollutants measured using different metrics can be evaluated for comparative analysis.

Note that for the formulation in Eqn. (2.5): 1) $MI(X_1, X_2)$ is the mutual information calculated based on partition of pairwise air pollution samples to $n_{X_1} \times n_{X_2}$, with which their joint and marginal probability distributions are estimated accordingly; 2) $N = S^{0.6}$ is an empirical parameter setting recommended by Reshef *et al* [88]; 3) theoretically, the base of logarithm (also for the above equations with logarithm operation) can be replaced with e, 10 or other values. Here, base 2 is used to evaluate the amount of information with unit bit. 4) the maximal value of $MI(X_1, X_2)$ with partitions $n_{X_1} \times n_{X_2}$ is $\log_2 \min(n_{X_1}, n_{X_2})$, and thus $MI(X_1, X_2)$ is normalized to $[0, 1]$ through the division operation; 5) in practice, calculating all $n_{X_1} \times n_{X_2}$ combinations is time-consuming. A sub-optimal $MIC(X_1, X_2)$ value is usually calculated instead; 6) $MI(X_1, X_2)$ is a general association metric between X_1 and X_2, while it does not provide information concerning the direction of correlation as PCC; 7) although the performance of MIC is arguable in comparison with MI [89], it provides a general metric for association measurement and facilitates the comparative analysis.

With the MIC of Eqn. (2.5), both the linear and nonlinear associations between pairwise air pollution variables can be captured and quantified in a general form based on information entropy. Moreover, this metric can also be used to identify the most significant predictor(s) for the health consequence of interest through evaluation

and sorting of their associations. In the following section, we will show the effect of MIC metric in correlation analysis and health predictors identification through experimental tests on the NAPS and NMMAPS datasets.

2.4 Experiments and Discussions

2.4.1 Tools for Experiments

With the above MIC formulation, quantitative experiments and analysis are made based on the aforementioned datasets. Several statistical tools are utilized to process the original data and facilitate the experimental process and the main tools are briefly introduced as follows.

i) The experiments are mainly performed using the Python (version 3.7.13) language in a virtual environment.

ii) `SciPy` (version 1.7.3) [90] is used for calculation of PCCs with Python. `SciPy` is a scientific computation library designed for efficient numerical computation based on `NumPy`, including various tools for applications in optimization, linear algebra, signal and image processing, etc. Its `scipy.stats` submodule for statistics is utilized for the following experiments.

iii) `minepy` (version 1.2.6) [91] is used for calculation of MICs with Python. `minepy` is a library developed for efficient computation of maximal information-related association. Its `minepy.MINE` submodule is utilized for computation of MICs in the following experiments.

2.4.2 Numerical Results

First, experiments on the NAPS datasets are performed to show the effect of PCC and MIC in association evaluation. Then MIC is used to identify the most significant predictor(s) in different health outcomes based on the NMMAPS datasets.

The association of different air pollutants (including temperature) evaluated using PCCs (absolute values) and MICs for Toronto, ON, Canada are presented as in the heat map of Figure 2.4. It can be seen from the figure that

a. The values of correlation measured using MICs are relatively milder than using PCCs in general.

b. Compared to PCCs, the influence of T on concentrations of air pollutants is smaller evaluated using MICs except for NO and NO_x. Both PCCs and MICs show that CO, NO, NO_2, NO_x and SO_2 are weakly associated with T, while MICs show a relatively stronger relationship between NO or NO_x and T. This implies that PCCs are not able to capture all the association and cannot fully reflect their relevance for predictors screening and variable selection.

c. CO and $PM_{2.5}$ are correlated with the other six elements evaluated using both PCCs and MICs, with MICs showing weaker relations. SO_2 and O_3 are also shown to have some relevance with other five pollutants using both PCCs and MICs. However, MIC shows a mild relationship between SO_2 and O_3, where their PCC shows they are almost independent. This is another case that PCC does not fully capture the association between different pollutants.

d. Although the MIC values are relatively smaller, both PCCs and MICs show strong

correlations of NO, NO_2, and NO_x, as these pollutants usually come from human activity and have the same sources like industrial emission, fossil fuel electric utilities and motor vehicles.

e. Compared to PCC, as MIC can capture multiple types of association instead of only linearity, it provides a more effective approach for dependency analysis and has the potential to facilitate predictor(s) selection for air pollution-related health assessment models like GAMs as well as analysis of their risk contributions therein.

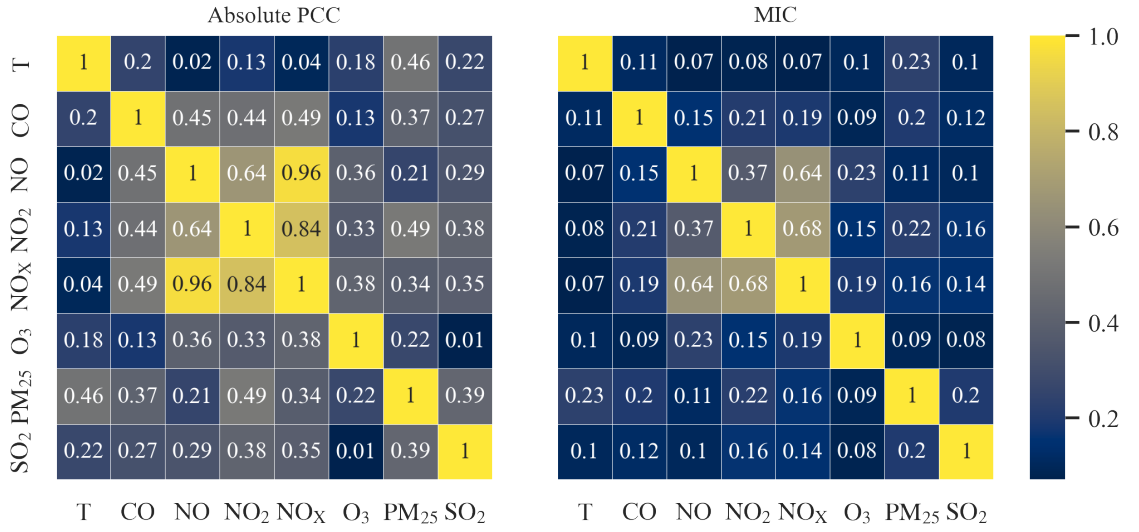

Figure 2.4: Comparison of PCCs and MICs with data of all seasons for Toronto, ON, Canada.

As the health risk in air pollutant-related risk assessment models like GAMs are usually evaluated for different seasons, relations of different air pollutants using data of both warm (from April to September) and cold (from October to March) seasons are evaluated and presented in Figure 2.5 and Figure 2.6, respectively. It can be seen from the figures that in both cold and warm seasons association evaluated

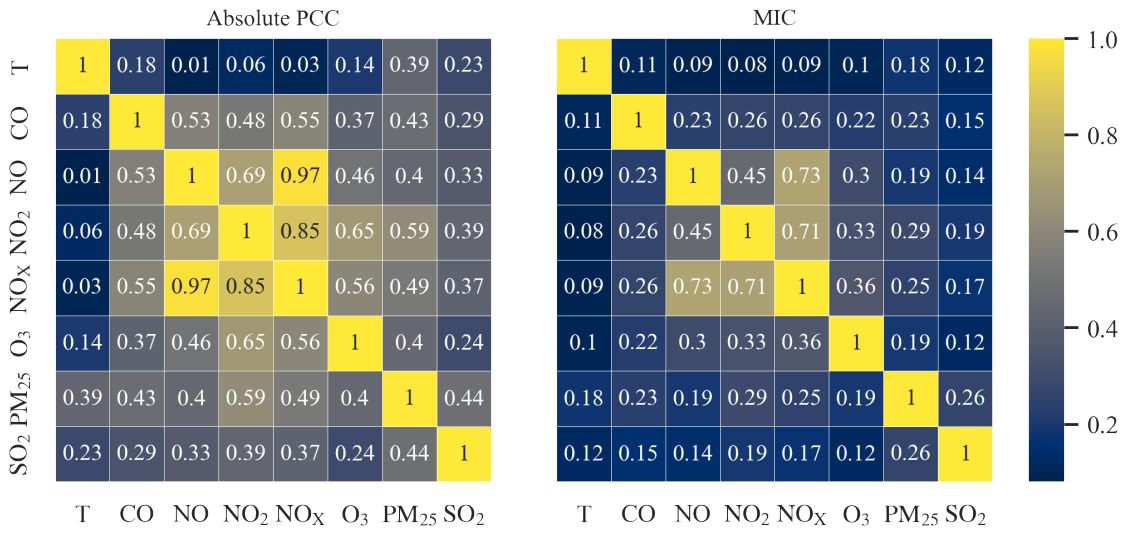

Figure 2.5: Comparison of PCCs and MICs with data of cold seasons (from October to March) for Toronto, ON, Canada.

using MICs are milder than using PCCs in general, following the patterns in Figure 2.4. However, MIC values for T and NO, T and NO_2, and T and NO_x in cold seasons, as well as MIC values for T and NO_x, CO and O_3, NO and $PM_{2.5}$, NO_2 and O_3 in warm seasons are all larger than their PCC values. This means that in these different season scenarios MIC still shows its capability to capture more

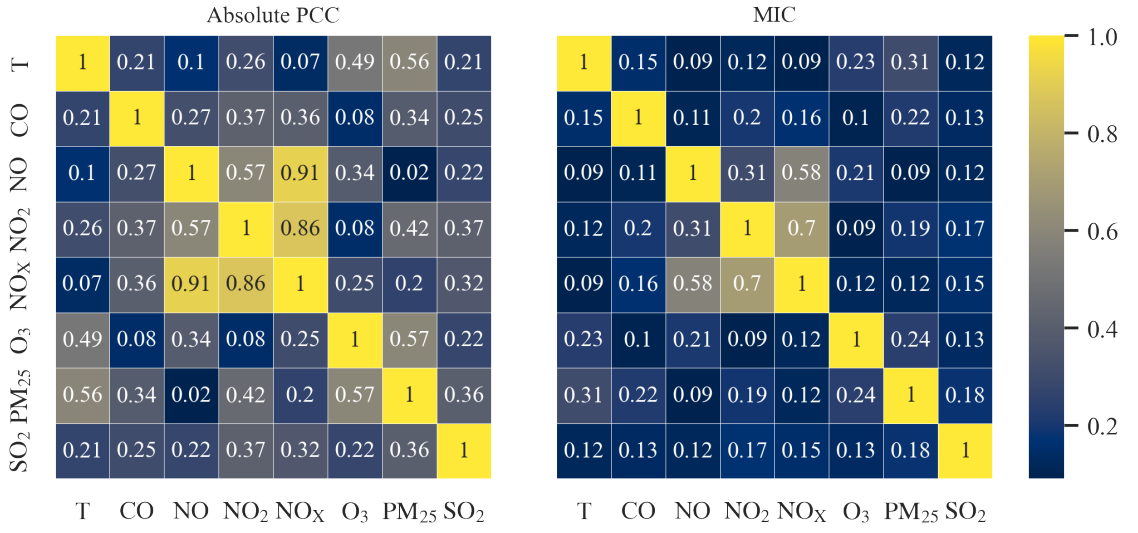

Figure 2.6: Comparison of PCCs and MICs with data of warm seasons (from April to September) for Toronto, ON, Canada.

association between different air pollutants compared to PCC, and that PCC may lead to incorrect association conclusions in nonlinear situations.

Note that in the above analysis: 1) values of PCC are mainly used for comparison and they might be meaningless in situations where the association of different elements is not linear; 2) values of MIC mean more for comparative analysis than for correlation strength identification; 3) MIC measures the general correlation without reporting correlation type or direction.

Experiments are further performed using datasets of Chicago, Illinois. MIC is used to evaluate the correlation between different predictors and mortality from various causes, in order to identify the most significant air pollution (and temperature) factors concerning the outcome. Test results are presented in Table 2.1 and it can be seen that:

a. Generally, older people (Age-3 category) are more sensitive to T than the other two categories for different health outcomes. SO_2 has relatively low association with all health outcomes of three age categories compared to other predictors. CO has relatively higher association with all health outcomes of three age categories than NO_2, O_3 and PM_{10} .

b. For the differing health outcomes of age categories Age-1 and Age-2, the MIC of PM_{10} is greater than that of both NO_2 and O_3. For the health outcomes PNEINF, PNEU, and RESP of Age-3, PM_{10} has also higher association compared with NO_2 and O_3, while for all-cause mortality, CVD and COPD of Age-3, O_3 has the maximum MIC of the three air pollutants.

Health Outcomes	Age Category	MIC					
		T	CO	NO_2	O_3	PM_{10}	SO_2
All-cause Mortality	Age-1	0.0637	0.0722	0.0389	0.0478	0.054	0.0314
	Age-2	0.0777	0.0736	0.0393	0.0418	0.0478	0.0347
	Age-3	0.1217	0.0729	0.0413	0.0681	0.0553	0.0273
	Total	0.1443	0.0748	0.0485	0.0624	0.0575	0.0321
CVD	Age-1	0.0634	0.0732	0.0431	0.0422	0.0503	0.0302
	Age-2	0.0686	0.0719	0.0397	0.0409	0.0504	0.0372
	Age-3	0.1161	0.0779	0.0416	0.0664	0.0524	0.0277
	Total	0.1186	0.0767	0.0409	0.0597	0.0595	0.0358
COPD	Age-1	0.0503	0.0695	0.0451	0.0421	0.0532	0.0268
	Age-2	0.0468	0.0746	0.039	0.0443	0.0498	0.0243
	Age-3	0.0579	0.0722	0.0415	0.0496	0.047	0.0267
	Total	0.0595	0.0681	0.0389	0.0443	0.0491	0.0271
PNEINF	Age-1	0.0549	0.0761	0.0382	0.0441	0.0551	0.0295
	Age-2	0.0588	0.0739	0.0389	0.0381	0.0495	0.0248
	Age-3	0.0861	0.0746	0.0406	0.0517	0.0565	0.0224
	Total	0.0937	0.0735	0.0412	0.0551	0.0553	0.023
PNEU	Age-1	0.0546	0.0763	0.0381	0.0432	0.055	0.0294
	Age-2	0.0582	0.0732	0.0387	0.038	0.0494	0.0248
	Age-3	0.0857	0.0746	0.0409	0.052	0.0562	0.0225
	Total	0.0923	0.0737	0.0408	0.0553	0.0551	0.0227
RESP	Age-1	0.0581	0.0708	0.0418	0.0477	0.0562	0.0252
	Age-2	0.055	0.076	0.0386	0.0442	0.0491	0.0226
	Age-3	0.0891	0.0767	0.0449	0.0561	0.057	0.027
	Total	0.0954	0.0672	0.0432	0.0581	0.0504	0.0284

Table 2.1: MICs between air pollutants and daily mortality from different causes for Chicago, IL, the U.S. All-cause Mortality is the all-cause non-accidental daily mortality. CVD, COPD, PNEINF, PNEU are RESP are short for cardiovascular deaths, chronic obstructive pulmonary disease, pneumonia and influenza, pneumonia and respiratory deaths. They represent the daily mortality from corresponding causes. Age-1 is for under 65 years of age, Age-2 is for 65 to 74, and Age-3 is for over 75.

c. For age category Age-1, CO has the greatest MIC on all health outcomes. With age increasing (from Age-2 to Age-3), T gradually becomes the most associated predictor for different health outcomes except COPD, where CO keeps the most associated one, although the influence of T increases and becomes the second most related factor.

d. For age category Total, the association between each health outcome and different air pollutants (temperature) is influenced by composition of the three age categories. All the health outcomes of Total are most associated with T and CO and least associated with SO_2 using the measurement ofMIC. Both O_3 and PM_{10} have greater MIC than NO_2 for different health outcomes.

e. Besides the above trend, for a specific health outcome, the most associated predictor(s) can be selected by sorting the corresponding MIC values in descending order.

For practical application in air pollution epidemiology, it is worth mentioning that: 1) MIC is a normalized form of MI that can evaluate the association between two different predictors (air pollutant exposure, temperature, etc.), or between a predictor and the health outcome for comparative analysis and main predictors identification; 2) MIC is a measurement between two random variables, which cannot decide if the impact of one predictor on health outcome of interest outweighs those of several other factors; 3) from both visual analysis of Figure 2.2 and the MIC values in Table 2.1, temperature (T) plays an important role in daily all-cause non-accidental mortality and mortality from several other diseases, especially for elder people,

and it might be worth pursuing further investigation of its role and formulation in GAM-based health outcome assessment models.

Chapter 3

LSTM Model for Health Outcome Assessment

In this chapter, an LSTM model is developed to evaluate the adverse impacts of air pollutants on public health. The LSTM network is used to extract dependencies from air pollution time series with distributed lags, and the dependent features are utilized for health outcomes assessment. Part of this chapter is published in the proceedings of the 6th International Conference on Statistics: Theory and Applications (ICSTA 2024) [92].

3.1 Introduction of LSTM Network

An LSTM network is a kind of artificial neural network proposed by Hochreiter and Schmidhuber [93] in 1997. It is an improved version of Recurrent Neural Network (RNN) [94, 95] that is developed to deal with sequential data and originates from

the recursive Hopfield network [96]. The elementary structure of RNN is shown in Figure 3.1, where $\boldsymbol{x}_{t-1}$, $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$ are sequential vector inputs, $\boldsymbol{h}_{t-1}$, $\boldsymbol{h}_t$ and $\boldsymbol{h}_{t+1}$ are sequential outputs, and the repeating module $\boldsymbol{A}$ is some neural network. Arrows represent the direction of information flow.

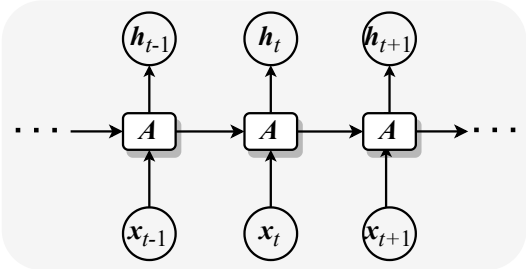

Figure 3.1: Elementary structure of RNN

One unique feature of the RNN in Figure 3.1 is that information from the previous processing module $\boldsymbol{A}$ can be utilized by its successor. The passage of "recurrent" information makes RNN the natural choice to deal with sequential data like voice or video sequences, as well as other kinds of time series data. However, the following several main disadvantages [97, 98] limit its application in practice:

1) Basic RNNs suffer from the problem of exploding or vanishing gradients. As the algorithm of backpropagation through time (BPTT) is usually applied for model training, then the error gradients are propagated and calculated based on the chain rule. Once the gradient of a processing module $\boldsymbol{A}$ becomes smaller or larger, multiplications of these gradients from $\boldsymbol{A}$s of different timestamps tend to explode or vanish, which makes RNNs difficult to train.

2) Basic RNNs can hardly process long-term sequential data. Influenced by the gradient issues, it is difficult for RNNs to capture the long-term dependencies in

longer sequences, where the training process becomes unstable.

3) Other important but more general issues met by most artificial neural networks like complexity of training, overfitting and generalization, depth of the networks, etc., also exist for traditional RNNs.

Aiming at the above issues of RNN, LSTM networks are specifically designed to mitigate the gradient exploding or vanishing problem and the long-term dependency problem. The key ideas of LSTM networks is to use a "cell state" to store information and use "three gates" to regulate the information flow. More specifically, the "cell state" stores information over time, where for each timestamp it interacts with the "three gates" for state update. The first step is the interaction with the "forget gate", which decides how much old information from previous timestamp needs to be removed from the "cell state". Then the "input gate" regulates how much new information from present timestamp is to be added to the "cell state". At last, through the "output gate", the "cell state" is optionally passed to the next time step after manipulation and update. A detailed mathematical description of the above process used in the air pollution-related health assessment problem is presented in Section 3.3 below.

With the gate mechanism regulating information flow and alleviating the gradient issues, LSTM network can store "short-term memory" in "long" time steps and capture the long-term dependency. Thus, it has been successfully applied to a large set of machine learning problems that need to handle the time-dependency issue across long time gaps of sequential data, such as natural language processing [99],

speech recognition [100], machine translation [101], and time series forecasting [102]. As to the air pollution-related health impacts problem, people are generally exposed to multiple air pollutants across long periods before some health events occur, where the air pollutant concentration levels are typical time series. This inspires the exploration of applying LSTM networks to capture the dependent impacts of previous air pollution exposure on future public health outcomes, i.e., distributed lags in this research.

3.2 Data Reorganization

In order to facilitate the following formulation and description of a LSTM network-based health outcomes assessment model, the air pollution time series data and health outcome data of Section 2.1 are reorganized as follows.

First, the original air pollutants data and health outcomes data are scaled through standardization:

$$x_{\mathrm{s},i} = \frac{x_i - \mu_x}{\sigma_x}, \tag{3.1}$$

where x_i is the i^{th} data point in the original time series datasets for each category, μ_x and σ_x are the mean and standard deviation values, and $x_{\mathrm{s},i}$ is the corresponding standardized value. After standardization, each variable of the original data is scaled to have a mean value of 0 and a standard deviation of 1. The main purposes of the above standardization operation are to:

1) reduce the influence of scales from different kinds of data, as ambient air pollu-

tants are generally measured and recorded with different metrics and scales;

2) accelerate the process of searching for the optimal network parameters during the training of the LSTM;

3) improve the generalization ability of the network by reducing overfitting.

Second, the datasets are further reorganized, with each element being a vector consisting of a series of the standardized data. Assuming the length of vectors after reorganization is m, the vectorized data sample can then be expressed as

$$\boldsymbol{x}_{\mathrm{s},i} = [x_{\mathrm{s},i-m+1}, x_{\mathrm{s},i-m+2}, \cdots, x_{\mathrm{s},i}], \tag{3.2}$$

where $\boldsymbol{x}_{\mathrm{s},i}$ is the i^{th} reorganized data sample. The new reorganized datasets are

$$[\boldsymbol{x}_{\mathrm{s},1}, \boldsymbol{x}_{\mathrm{s},2}, \cdots, \boldsymbol{x}_{\mathrm{s},n-m+1}], \tag{3.3}$$

where n is the length of original data series, becoming $n - m + 1$ after reorganization. The purpose of this operation is to facilitate the inputs during training of the following LSTM model. It is worth mentioning that parameter m has the same meaning as the time lag used in GAM-type health assessment models and the expression in Eqn. (3.2) includes distributed lags which means sequential exposures from timestamp $i - m + 1$ to timestamp i.

3.3 LSTM Model for Health Outcome Assessment

3.3.1 LSTM Network for Feature Information Extraction

For some public health outcome of interest at the time period t, the LSTM network used for sequential information extraction from multiple air pollution time series

of length m, i.e., exposure from previous m time periods, is shown in Figure 3.2. The LSTM network has r stacked layers to process the input sequence and each layer has m repeating neural network modules $\boldsymbol{A}$. The input to the LSTM model is $[\boldsymbol{x}_{\mathrm{s},t-m+1}, \boldsymbol{x}_{\mathrm{s},t-m+2}, \cdots, \boldsymbol{x}_{\mathrm{s},t}]$, i.e., multiple air pollutant concentration (exposure) sequence of previous m periods, and its output is the extracted feature vector sequence with temporal dependency $[\boldsymbol{h}_{r,t-m+1}, \boldsymbol{h}_{r,t-m+2}, \cdots, \boldsymbol{h}_{r,t}]$.

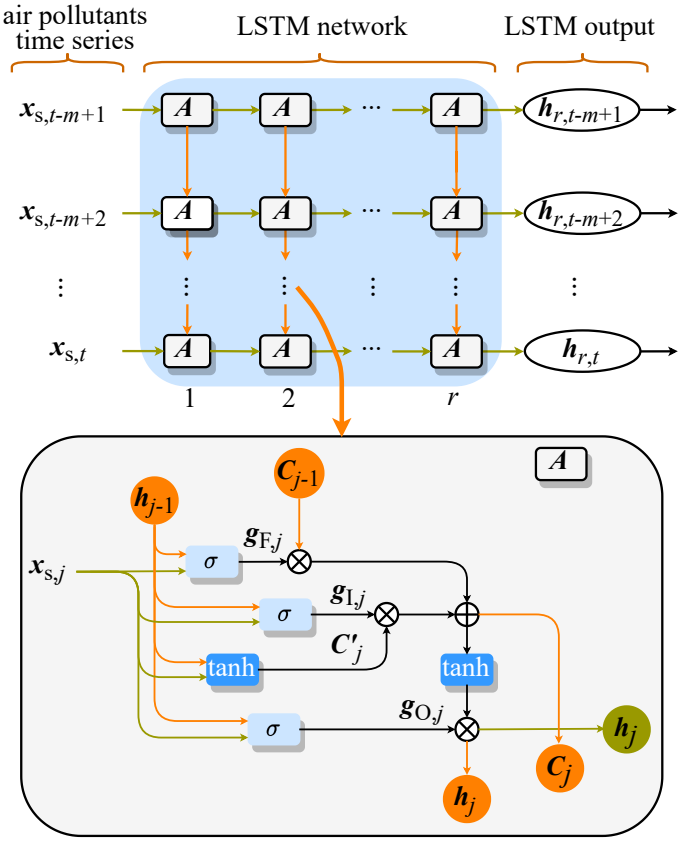

Figure 3.2: LSTM network for sequential information extraction.

Note that in Figure 3.2: 1) each element $\boldsymbol{x}_{\mathrm{s},j}$ $(j = t - m + 1, \cdots, t.)$ of the input air pollution sequence is a vector consisting of multiple air pollutants, which is a generalized expression of the scalar element in Eqn. (3.2) for single air pollutant; 2) the dimension of the output of module $\boldsymbol{A}$ for each layer, $\boldsymbol{h}_{k,j}$ $(k = 1, \cdots, r,\ j =$

$t-m+1, \cdots, t.$) is l, which can be different from the dimension of the input, $\boldsymbol{x}_{\mathrm{s},j}$; and 3) m, r and l are three parameters that determine the network structure.

The repeating module $\boldsymbol{A}$ uses three specifically designed "gates" and a "cell state" to regulate and extract information from the input, where "gates" are essentially sigmoid functions (between 0 and 1) that decide the information flow, with 0 for "completely blocked" and 1 for "completely passed", and "cell state" is used to store the dependent information extracted from the input sequence. As operations for different layers are alike, layer number k is neglected in the following description.

For period j, how much information from "cell state" $\boldsymbol{C}_{j-1}$, i.e., stored information from last period $j-1$, should be blocked is decided by $\boldsymbol{g}_{\mathrm{F},j}$ of the "forget gate" as

$$\boldsymbol{g}_{\mathrm{F},j} = \sigma(\boldsymbol{w}_{\mathrm{F}} \cdot \boldsymbol{x}_{\mathrm{s},j} + \boldsymbol{u}_{\mathrm{F}} \cdot \boldsymbol{h}_{j-1} + \boldsymbol{b}_{\mathrm{F}}), \tag{3.4}$$

where $\boldsymbol{w}_{\mathrm{F}}$, $\boldsymbol{u}_{\mathrm{F}}$ and $\boldsymbol{b}_{\mathrm{F}}$ are network weights and biases. $\sigma(\cdot)$ is a sigmoid function. The value of $\boldsymbol{g}_{\mathrm{F},j}$ is decided by both the air pollution concentration $\boldsymbol{x}_{\mathrm{s},j}$ in the j^{th} time period and extracted information $\boldsymbol{h}_{j-1}$ from the $(j-1)^{\mathrm{st}}$ time period.

Then the "input gate", Eqn. (3.5), is used to decide how much new information from the candidate "cell state", Eqn. (3.6), could be added to $\boldsymbol{C}_{j-1}$:

$$\boldsymbol{g}_{\mathrm{I},j} = \sigma(\boldsymbol{w}_{\mathrm{I}} \cdot \boldsymbol{x}_{\mathrm{s},j} + \boldsymbol{u}_{\mathrm{I}} \cdot \boldsymbol{h}_{j-1} + \boldsymbol{b}_{\mathrm{I}}) \tag{3.5}$$

$$\boldsymbol{C}'_{j} = \tanh(\boldsymbol{w}_{\mathrm{C}} \cdot \boldsymbol{x}_{\mathrm{s},j} + \boldsymbol{u}_{\mathrm{C}} \cdot \boldsymbol{h}_{j-1} + \boldsymbol{b}_{\mathrm{C}}), \tag{3.6}$$

where $\boldsymbol{w}_{\mathrm{I}}$, $\boldsymbol{u}_{\mathrm{I}}$, $\boldsymbol{w}_{\mathrm{C}}$, and $\boldsymbol{u}_{\mathrm{C}}$ are network weights, and $\boldsymbol{b}_{\mathrm{I}}$, $\boldsymbol{b}_{\mathrm{C}}$ are network biases. Note that tanh is the hyperbolic tangent activation function. Together with the "forget

gate", the "cell state" output from the j^{th} module is updated as $\boldsymbol{C}_j$ in Eqn. (3.7):

$$\boldsymbol{C}_j = \boldsymbol{g}_{\text{F},j} \odot \boldsymbol{C}_{j-1} + \boldsymbol{g}_{\text{I},j} \odot \boldsymbol{C}'_j, \tag{3.7}$$

where $\odot$ is the Hadamard product. The new "cell state" $\boldsymbol{C}_j$ blocks part of old information $\boldsymbol{C}_{j-1}$ from last time period $j-1$ and adds part of new information $\boldsymbol{C}'_j$ from current time period. Finally, $\boldsymbol{C}_j$ is passed to the next time period $j+1$ as well as to the next network layer after being processed by the "output gate" as:

$$\boldsymbol{g}_{\text{O},j} = \sigma(\boldsymbol{w}_{\text{O}}\boldsymbol{x}_{\text{s},j} + \boldsymbol{u}_{\text{O}}\boldsymbol{h}_{j-1} + \boldsymbol{b}_{\text{O}}) \tag{3.8}$$

$$\boldsymbol{h}_j = \boldsymbol{g}_{\text{O},j} \odot \tanh(\boldsymbol{C}_j), \tag{3.9}$$

where $\boldsymbol{w}_{\text{O}}$, $\boldsymbol{u}_{\text{O}}$ and $\boldsymbol{b}_{\text{O}}$ are weights and biases. The output $\boldsymbol{h}_j$ of a module $\boldsymbol{A}$ is decided by both the updated "cell state" $\boldsymbol{C}_j$ and the "output gate" $\boldsymbol{g}_{\text{O},j}$.

With the above regulation through m time steps of r layers, dependent information $\boldsymbol{h}_{r,j}$ $(j = t-m+1, t-m+2, \cdots, t)$ is extracted from the input air pollution sequence, which can be further used to assess the air pollution-related health outcome of interest. Note that in the above description: 1) the inputs to each layer of LSTM network are outputs of module $\boldsymbol{A}$s from the previous layer except for the first one where its inputs are the air pollution sequences; 2) for the first time period of each layer (j=1), its input $\boldsymbol{h}_{j-1}$ and $\boldsymbol{C}_{j-1}$ from last time step are zeros; and 3) the meaning of extracted feature vector $\boldsymbol{h}_{r,j}$ $(j = t-m+1, t-m+2, \cdots, t)$ is determined by the health outcome of interest and may be different for various health events.

3.3.2 Health Outcome Assessment

With the sequential information $[\boldsymbol{h}_{r,t-m+1}, \boldsymbol{h}_{r,t-m+2}, \cdots, \boldsymbol{h}_{r,t}]$ extracted from the input air pollution sequence through the LSTM network, health outcomes of interest are assessed using combinations of these features. As these extracted features from previous exposure of m time periods may have different contributions to the health outcome at time period t, weighed evaluation of the lagged effects is used for health outcome assessment as shown in Figure (3.3).

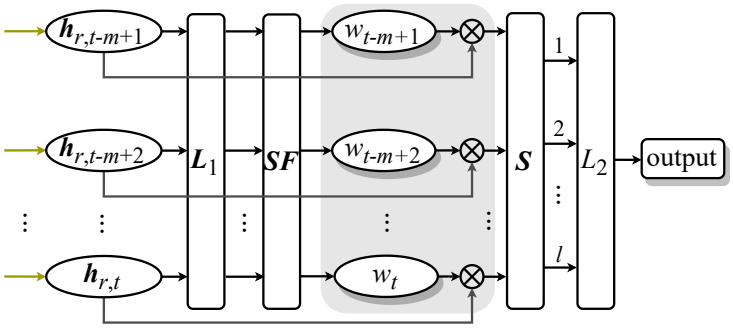

Figure 3.3: Health outcome assessment with weighted evaluation of lags

First, a linear transformation layer $\boldsymbol{L}_1$ with m processing units is used to combine the l features in each $\boldsymbol{h}_{r,j}, j = t-m+1, t-m+2, \cdots, t$,

$$\boldsymbol{L}_1(j) = \sum_{i=1}^{l} w_{i,j} \boldsymbol{h}_{r,j}(i) + b_j, j = t-m+1, t-m+2, \cdots, t, \tag{3.10}$$

with $w_{i,j}$ and b_j the weight and bias, respectively. Then, a softmax layer $\boldsymbol{SF}$ is used to generate weights for the output sequence of $\boldsymbol{L}_1$ by mapping their values to $(0, 1)$:

$$w_j = \boldsymbol{SF}(j) = \frac{e^{\boldsymbol{L}_1(j)}}{\sum_{j=t-m+1}^{t} e^{\boldsymbol{L}_1(j)}}, j = t-m+1, t-m+2, \cdots, t, \tag{3.11}$$

where w_j is the weight for the influence of air pollution exposure in time period j on some health outcome. After that, a sum layer $\boldsymbol{S}$ is used to combine each feature of

$\boldsymbol{h}_{r,j}$ from m time periods with weighted evaluation of their impacts as Eqn. (3.12):

$$\boldsymbol{S}(i) = \sum_{j=t-m+1}^{t} w_j \boldsymbol{h}_{r,j}(i), i = 1, 2, \cdots, l. \tag{3.12}$$

Finally, the impacts of l features are combined to assess the health outcome of interest through a linear transformation layer L_2:

$$\text{output} = L_2[\boldsymbol{S}(i), i = 1, 2, \cdots, l] = \sum_{i=1}^{l} w_i' \boldsymbol{S}(i) + b, \tag{3.13}$$

where w_i' and b are weight and bias.

Note that in the above formulation: 1) exposure to air pollution from $t-m+1$ to t is used to assess the health outcome at t, where other periods of interest can be used instead; 2) w_j in Eqns. (3.11) and (3.12) is the weight for the j^{th} feature vector $\boldsymbol{h}_{r,j}$ and is shared by all l features therein to evaluate the impacts of distributed previous exposures; 3) w_i' in Eqn. (3.13) is the weight evaluating combined impacts of the i^{th} element in all the extracted feature vectors; 4) the LSTM network is used to process and extract chronically dependent information from the input air pollutant series through each layer, as well as to capture and fit the relation between input sequence and health outcome through multiple layers together with the subsequent evaluation layers; 5) the output from L_2 is the health outcome of interest at time period t, which can be mortality count, morbidity count, cardiovascular disease count, and other metrics.

3.4 Results and Discussions

Performance of the proposed model was tested through experiments on datasets from Chicago and the `NMMAPS` database. Daily all-cause non-accidental mortality for three age groups were used for assessment. Besides T, air pollutants of interest were selected from NO_2, CO, SO_2, O_3 and PM_{10}. All categories of data were preprocessed, reorganized and split into two parts, with 70% for model training and the remaining 30% for performance evaluation. Note that due to the temporal structure of the data, the split was enforced to give contiguous data; that is, the train was the first 70% and the test was the final 30%. The following three experiments were performed.

i) For different lengths of input air pollutants sequence m, experiments were performed to demonstrate the performance in capturing the impacts of distributed lags (previous exposure).

ii) Single-pollutant and multipollutant models are constructed and tested to show the performance in adapting to both single and multiple air pollutant (s).

iii) Comparison was made with a standardized GAM model widely used in air pollution epidemiology, Eqn. (3.14), to illustrate the advantage of proposed model,

$$\log(\mu_t) = \beta_0 + \beta_1\chi_t + \beta_2 \text{DOW}_t + \text{NS}(T, df = 3)_t + \text{NS}(\text{Time,df} = 6/\text{yr})_t \quad (3.14)$$

where μ_t is the daily mortality, χ_t is the daily mean concentration level of air pollutant of interest, β_0, β_1 and β_2 are linear coefficients, DOW_t is a factor variable for day of week, and NS is the natural cubic spline regression function with degree of freedom $df = 3$ for daily mean temperature T and $df = 6$ per year for *Time*. Note that to accomplish this, the data that was fed into the LSTM

was pre-filtered to account for time, so allow for a more accurate comparison.

3.4.1 Performance Metrics

The performance metrics used to evaluate all experiments are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Explained Variation (R^2). Their formulas are shown in Eqns. (3.15 - 3.18).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{\mathrm{P},i} - y_{\mathrm{T},i})^2} \tag{3.15}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_{\mathrm{P},i} - y_{\mathrm{T},i}| \tag{3.16}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_{\mathrm{T},i} - y_{\mathrm{P},i}}{y_{\mathrm{T},i}}\right| \tag{3.17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_{\mathrm{T},i} - y_{\mathrm{P},i})^2}{\sum_{i=1}^{n}(y_{\mathrm{T},i} - \bar{y}_{\mathrm{T}})^2} \tag{3.18}$$

where $y_{\mathrm{P},i}$ and $y_{\mathrm{T},i}$ are the i^{th} predicted daily mortality and its actual value, $\bar{y}_{\mathrm{T}}$ is the mean actual value, and n is the total number of samples used for evaluation. Both performance metrics are calculated using the standardized data in all experiments.

3.4.2 Parameter Settings

`PyTorch` (version `1.8.1+cpu`) was used to construct the LSTM model. The main hyperparameters of the LSTM network during training and test are listed in Table 3.1. The optimizer used was the Adaptive Moment Estimation (ADAM) algorithm, where

a variable learning rate with a decay rate of 0.95 per 100 epochs was used to accelerate the convergence of model training process.

Parameter	Value	Description
r	3	number of hidden layers
m	1–12	length of input sequence
l	13	dimension of extracted feature
epochs	5000	training time periods
lr	5e-2	initial learning rate
decay	0.95	decay of *lr* every 100 epochs
seed	65	random seed

Table 3.1: Main parameters setting for the LSTM model.

3.4.3 Numerical Results

Tests with different m

For T, PM_{10} and O_3, the losses during the training process, performance metrics on both training and test sets, and comparisons between predicted mortality and actual count for different lengths of input air pollution sequence m are shown in Fig. 3.4, Table 3.2 and Fig. 3.5. The variation in the losses in Fig. 3.4 demonstrates that for different lengths of input air pollution sequence the proposed LSTM model has good convergence properties, and lagged impacts of distributed exposure can be accommodated for daily mortality assessment. Performance metrics in Table 3.2 show that both the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values have general downward trends on the training set with the length of input exposure sequence, i.e., m increasing. With a longer term of exposure

sequence, more useful features and information can be extracted through the LSTM network for mortality assessment. For issues of data noises, distribution differences and possible overfitting, downward trends of the performance metrics are not obvious on the test set.

Another two experiments were performed to further demonstrate the performance of the proposed model with different lengths of air pollution sequence m. Besides T, the air pollutants were selected based on the MIC between each pollutant and the mortality, and the most associated three and four air pollutants, i.e., $PM_{10} + O_3 + CO$ and $PM_{10} + O_3 + CO + NO_2$ were used for tests. Experimental results are shown in Table 3.3 and Fig. 3.6 through Fig. 3.9. Similar to Fig. (3.4), the training losses in Figs. 3.6 and 3.8 further demonstrate the good convergence properties of the model, and comparisons in Figs. 3.7 and 3.9 roughly show that the LSTM model can capture short-term variations and long-term trends in both tests.

	RMSE		MAE			RMSE		MAE	
	TrS	TeS	TrS	TeS		TrS	TeS	TrS	TeS
m=1	0.614	1.357	0.441	1.040	m=7	0.080	1.459	0.047	1.156
m=2	0.192	1.909	0.098	1.455	m=8	0.048	1.387	0.030	1.114
m=3	0.095	1.679	0.042	1.323	m=9	0.055	1.460	0.032	1.157
m=4	0.094	1.574	0.047	1.219	m=10	0.066	1.394	0.034	1.108
m=5	0.092	1.561	0.050	1.223	m=11	0.059	1.398	0.031	1.127
m=6	0.042	1.472	0.026	1.145	m=12	0.049	1.372	0.028	1.101

Table 3.2: Performance comparison for different m on the training and test sets (TrS: training set, TeS: test set).

Downward trends are not obvious for both performance metrics in Table 3.3 with m increasing.

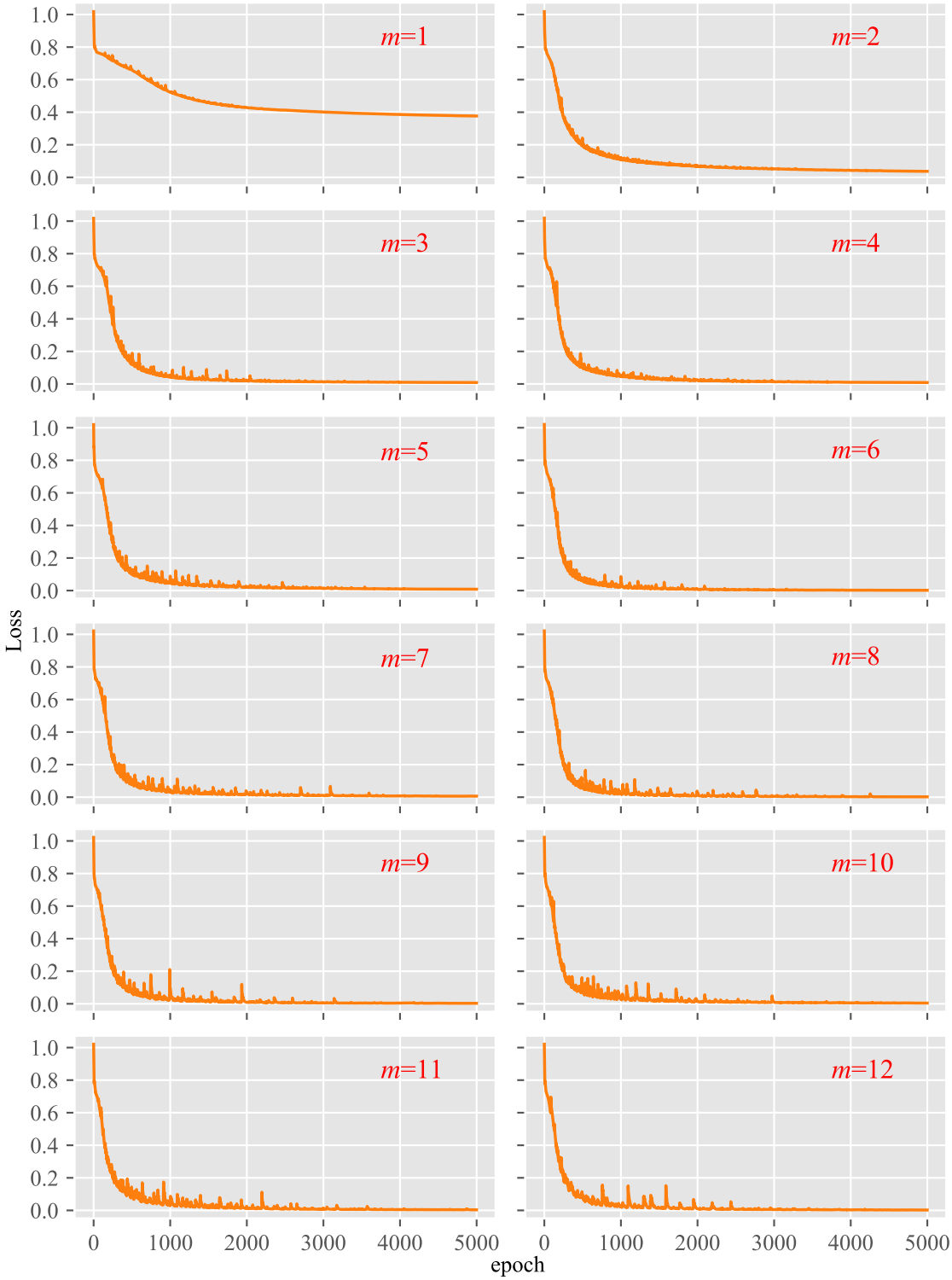

Figure 3.4: Training losses for different lengths of air pollution exposure sequence (m=1–12).

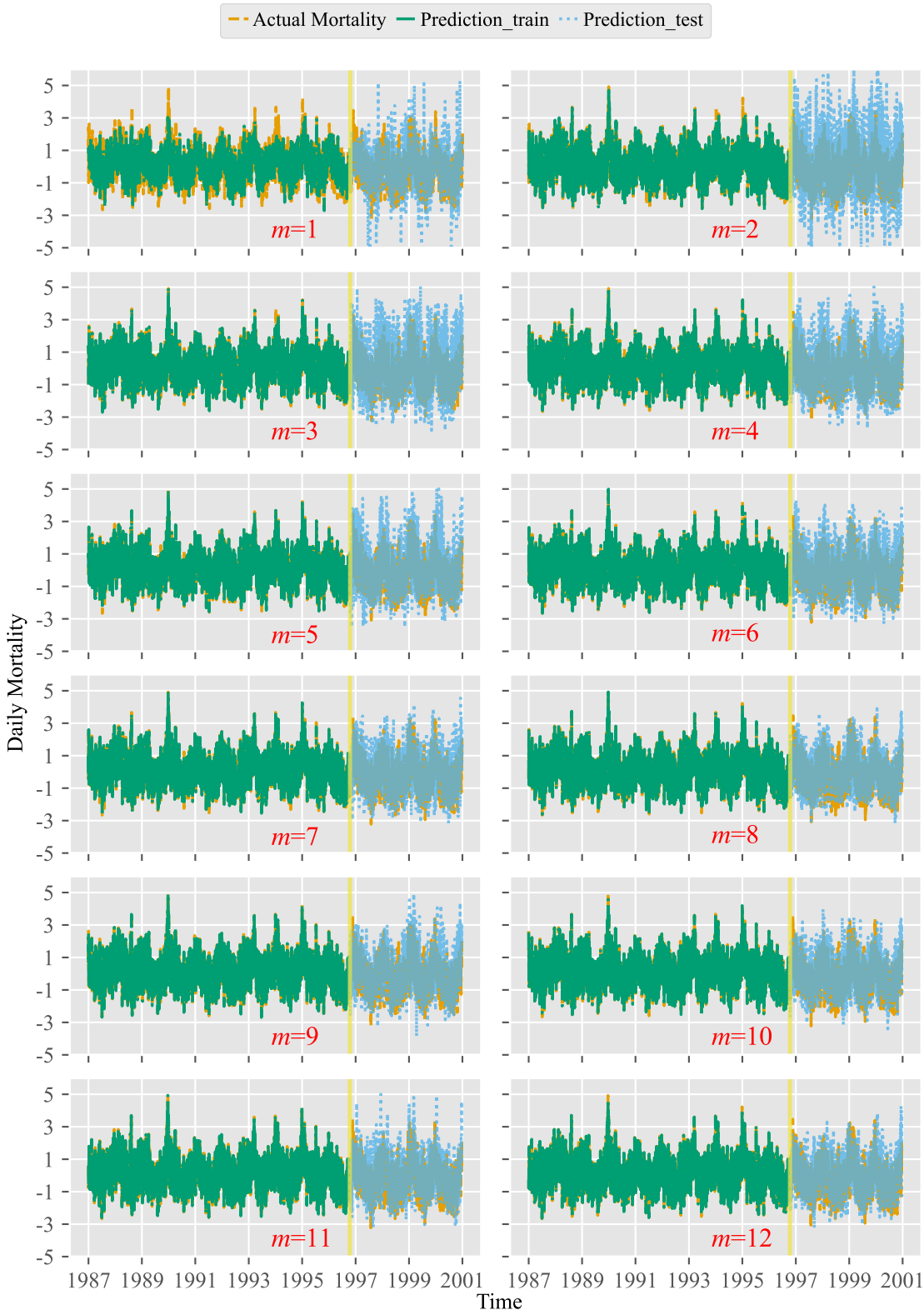

Figure 3.5: Comparison between prediction and actual daily mortality (m=1–12).

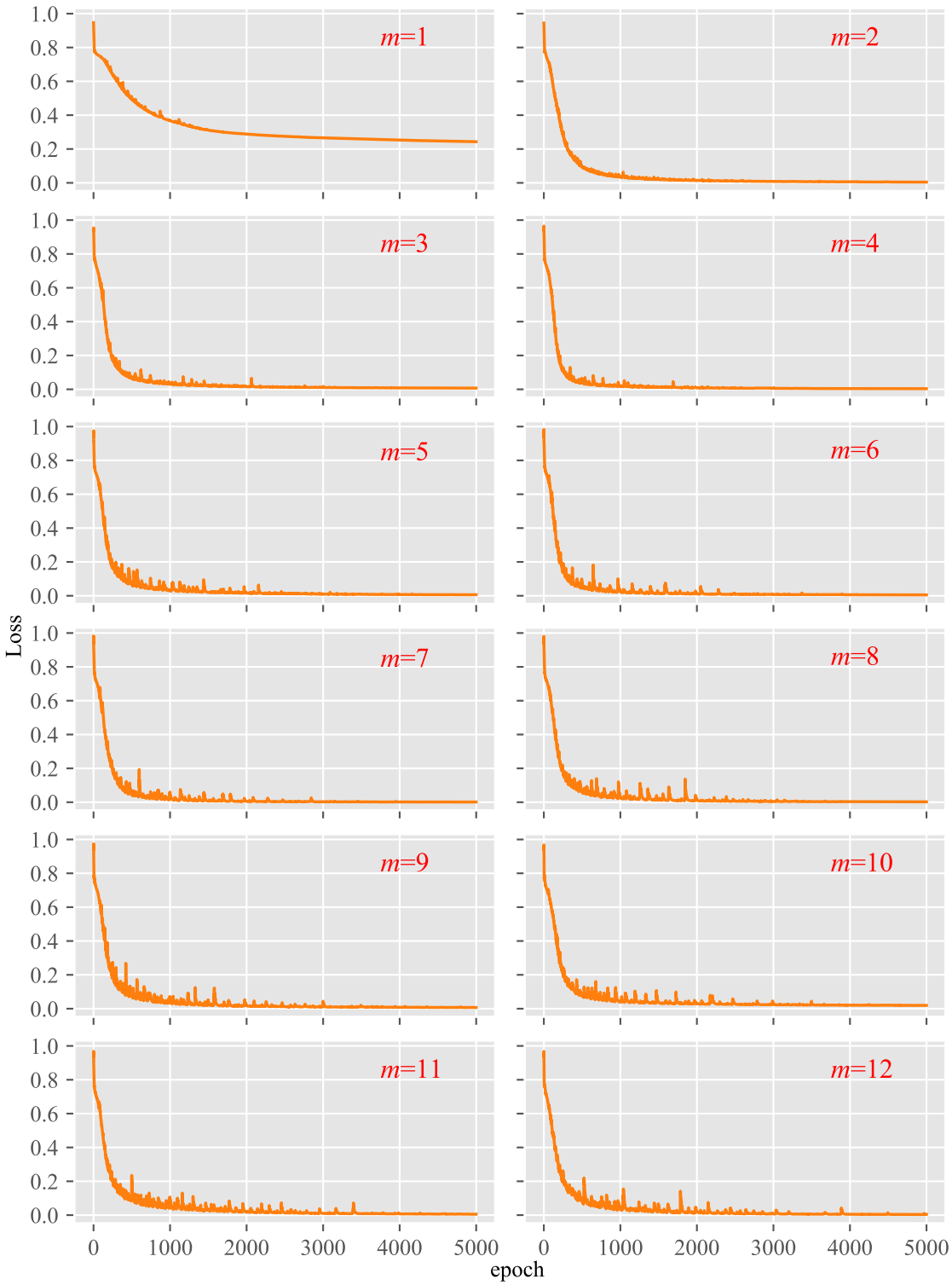

Figure 3.6: Training losses with air pollutants $PM_{10} + O_3 + CO$.

Figure 3.7: Comparison between prediction and actual daily mortality with air pollutants $PM_{10} + O_3 + CO$.

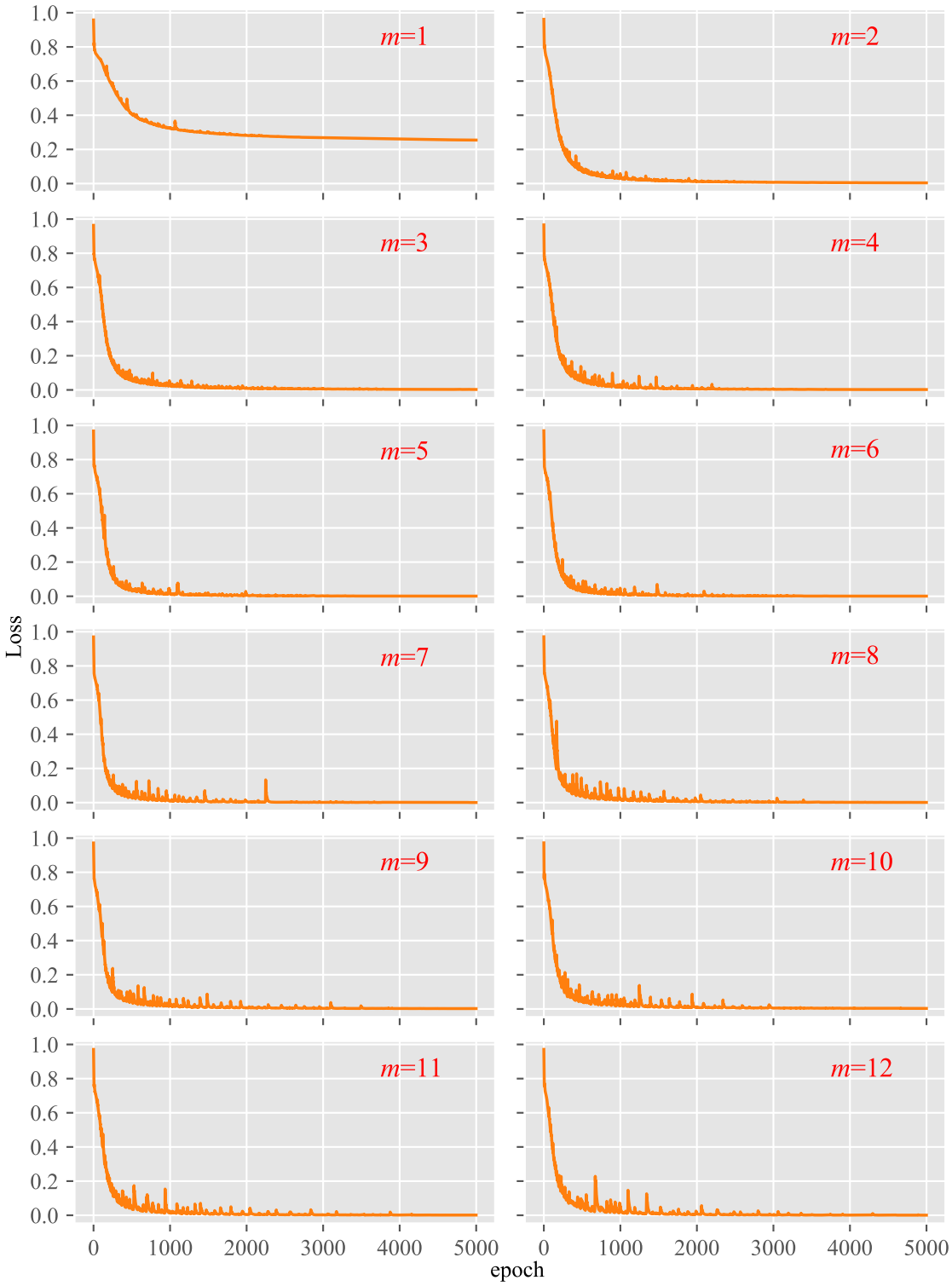

Figure 3.8: Training losses with air pollutants $PM_{10} + O_3 + CO + NO_2$.

Figure 3.9: Comparison between prediction and actual daily mortality with air pollutants $PM_{10} + O_3 + CO + NO_2$.

	$PM_{10} + O_3 + CO$					$PM_{10} + O_3 + CO + NO_2$			
	RMSE		MAE			RMSE		MAE	
	TrS	TeS	TrS	TeS		TrS	TeS	TrS	TeS
m=1	0.493	1.480	0.360	1.146	m=1	0.505	1.475	0.347	1.123
m=2	0.067	1.887	0.038	1.439	m=2	0.066	1.665	0.038	1.318
m=3	0.084	1.551	0.039	1.230	m=3	0.051	1.725	0.028	1.358
m=4	0.060	1.539	0.026	1.240	m=4	0.046	1.559	0.028	1.225
m=5	0.081	1.493	0.041	1.190	m=5	0.030	1.490	0.019	1.192
m=6	0.073	1.428	0.029	1.125	m=6	0.029	1.455	0.021	1.156
m=7	0.035	1.396	0.022	1.112	m=7	0.023	1.487	0.014	1.202
m=8	0.049	1.404	0.029	1.107	m=8	0.034	1.368	0.022	1.099
m=9	0.082	1.433	0.035	1.144	m=9	0.048	1.386	0.026	1.114
m=10	0.138	1.500	0.053	1.179	m=10	0.056	1.470	0.028	1.169
m=11	0.074	1.471	0.037	1.154	m=11	0.037	1.320	0.020	1.053
m=12	0.067	1.389	0.039	1.120	m=12	0.027	1.397	0.016	1.117

Table 3.3: Performance comparison with different air pollutants (TrS: training set, TeS: test set).

Tests for multiple air pollutants

Test results for different combinations of air pollutants and T are shown in Table 3.4 and Fig. 3.10. As shown in Fig. 3.10, the proposed model can accommodate multiple air pollutants and have good convergence properties for different combinations. As shown in Table 3.4, both RMSE and MAE have roughly downward trends on the training set with the kinds of air pollutants increasing, where more types of air pollutants can provide more mortality-related information to the LSTM network for feature extraction and outcome assessment. Both metrics on the test set are not as perfect as their counterparts on the training set. The first reason is that data distribution of the test set may not be exactly the same as that of the training

set, which lowers the model's generalization ability. The second one is that noise from air pollutant series of the training set may lead to overfitting that lowers the models' performance on the training set. Nonetheless, the proposed model shows great potential in multipollutant-related health outcome assessments as shown by comparison between prediction and actual mortality on the right-side subfigures in Fig. 3.10.

single/multiple air pollutant (s)	RMSE		MAE	
	TrS	TeS	TrS	TeS
PM_{10}	0.079	1.494	0.044	1.87
$PM_{10} + O_3$	0.081	1.460	0.047	1.157
$PM_{10} + O_3 + CO$	0.063	1.383	0.027	1.096
$PM_{10} + O_3 + CO + NO_2$	0.042	1.393	0.021	1.113
$PM_{10} + O_3 + CO + NO_2 + SO_2$	0.028	1.482	0.011	1.174

Table 3.4: Comparison of performance metrics with different air pollutant (s) (TrS: training set, TeS: test set, m=7).

Comparison with GAM

Performance of the LSTM-based model is compared with a standard GAM used in air pollution epidemiology as Eqn. (3.14) and the results are shown in Table 3.5 and Fig. 3.11. `R` (version 4.3.1) was used for the experiments of the GAM with the packages `mgcv` (version 1.8.42) and `splines` (version 4.3.1). The selected air pollutant of interest is PM_{10}, which has attracted lots of attention from researchers in this field. Distributed exposure sequence from period $t-m+1$ to t is input to the LSTM model for daily mortality assessment at period t and single exposure at $t-lag$ is used by the GAM for assessment. These two formulations are a distributed-lag

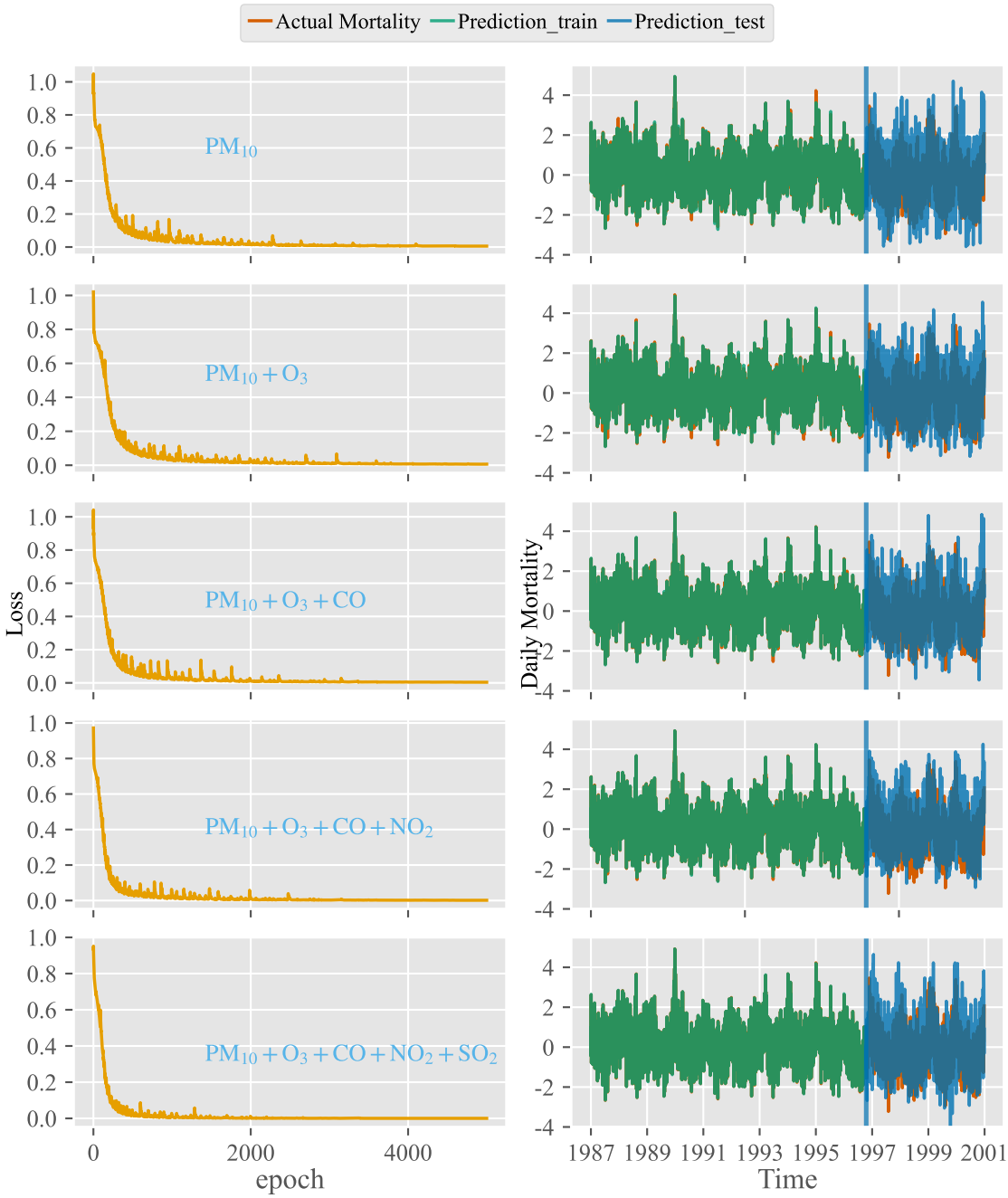

Figure 3.10: Training losses (left side) and comparison between prediction and actual mortality (right side) for single and multiple air pollutant (s) (m=7).

model and a single-lag model. In addition, as GAMs used in air pollution-related health risk assessment are generally for retrospective studies, all data are used to calculate the model parameters. As shown in Table 3.5, for all four experiments

both the RMSE and MAE of the LSTM model on the training set outperform their counterparts of the GAM. This demonstrates that the proposed LSTM model can better fit the health outcome, i.e., daily mortality using distributed lags, than the GAM using only a single lag. It also shows that in a retrospective study, the LSTM model can better assess the health outcome taking advantage of the impacts from distributed lags. Due to problems of data noise, data distribution differences and model overfitting as mentioned in previous experiments, performance of the LSTM model on the test set is not as good as on the training set and are exceeded by their counterparts of the GAM as showed in Table 3.5. However, both the RMSE and MAE of the GAM are calculated based on the training set, i.e., the whole dataset, and their counterparts of LSTM are calculated on the test set only, making direct comparisons more difficult.

In addition, there are several points worth mentioning for this comparative experiment:

1. GAMs applied in air pollution epidemiology are generally used to fit historical data and lack the capability of predicting future health outcome with new air pollution exposure data, i.e., they are used for retrospective studies, not for predicting;

2. limited by their generalized linear formulation, GAMs are more suitable for evaluating the impacts of single lag and single pollutant instead of distributed lags and multiple pollutants, which are the desired modeling outcomes (that is, they are not entirely fit for purpose);

3. the LSTM model has better adaptability in accommodating distributed lags and multiple air pollutants and can synthetically evaluate the impacts from historical air pollution exposure.

	LSTM					GAM	
	RMSE		MAE			RMSE	MAE
	TrS	TeS	TrS	TeS		TrS	TrS
m=2	0.2360	1.9299	0.1354	1.4055	lag=1	0.8818	0.6992
m=4	0.0961	1.6528	0.0519	1.2745	lag=3	0.8823	0.6994
m=6	0.1495	1.4961	0.0587	1.1822	lag=5	0.8819	0.6992
m=8	0.0744	1.4096	0.0363	1.1193	lag=7	0.8824	0.6994

Table 3.5: Comparison performance metrics between the LSTM model and the GAM (TrS: training set, TeS: test set).

Note that in the above numerical experimentation: 1) parameters of the LSTM model can be further tuned for performance improvement; 2) the downward trends for both metrics may not remain with m or types of air pollutants continue increasing due to redundant information brought by associated elements; 3) either the exposure periods of interest or combinations of air pollutants of interest can be adjusted for specific applications; 4) other health outcomes like morbidity or mortality from air pollutants of interest can also be evaluated instead.

3.4.4 Conclusion

At their most simplistic, we can think of the LSTM models as being nonlinear autoregressive models: similar to time series regressions, but with the relaxation of the linearity of the predictors, and similarly the relaxation of the explicit causal autoco-

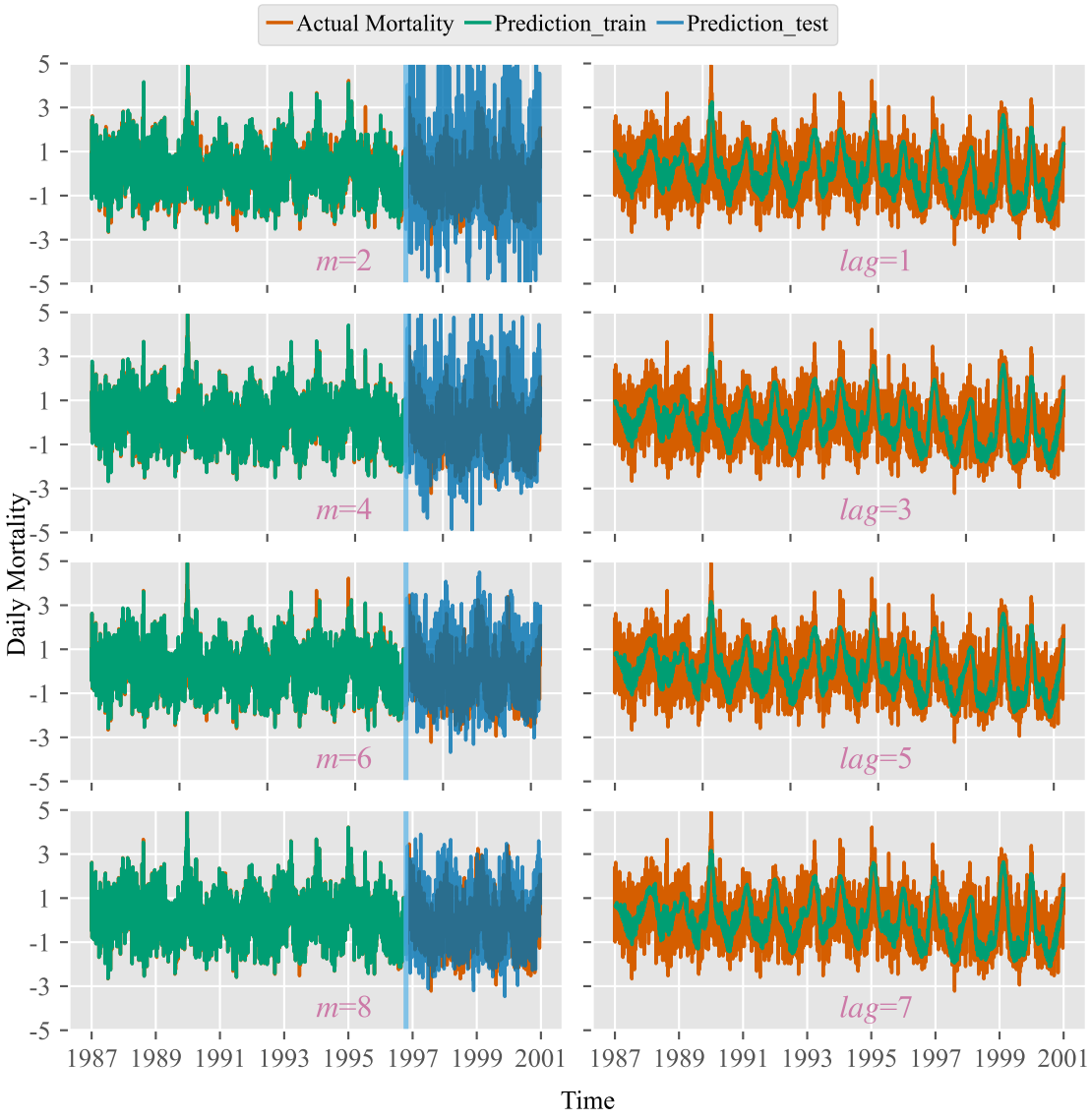

Figure 3.11: Comparison between the LSTM model and the GAM for daily mortality assessment.

variance structure. In this chapter, we have demonstrated that it is possible to fit and tune LSTM models on time-lagged air pollution and health outcome data and to obtain high precision predictive models.

The challenge which remains is how to take these predictive models and practically use them! In practice, the use of GAMs comes down to extracting one or more coefficients to particular predictors, treating these as "risks", and assuming that

they correspond to directly interpretable measurements of the potential impact (on human health) of rises in air pollution. This interpretation is lacking almost entirely from these LSTM models – they are predictive, but without the direct coefficients that we are used to. Typical LSTM models may have tens or hundreds of thousands of "coefficients", and we can't do much with those.

How to deal with this while retaining the incredible predictive power of LSTMs remains an open question, and in the next chapter we will begin exploring what we can, and cannot, say about these models, and the implications for the explanation of health outcomes that can be associated with air pollution exposure.

Chapter 4

Interpretability of Mortality-Estimate LSTM Model

To explore the reliability and utility of the LSTM model for the evaluation of health outcomes developed in Chapter Chapter 3, this chapter addresses the interpretability of the proposed model. Following a review of recent literature on interpretability techniques applied to LSTM models, we employ the LIME method to analyze the importance of characteristics within the proposed LSTM framework to assess pollution-health associations. Using LIME, we examine the temporal significance of predictors in estimating mortality outcomes based on climate and environmental data. Furthermore, this chapter explores the relationship between the importance of features, specifically, O_3 and PM_{10}, and temperature within the context of the LSTM model.

4.1 Explainable LSTM models

Although deep learning methods have demonstrated remarkable success in regression tasks, consistently excelling in performance metrics such as RMSE, MAE, Mean Absolute Percentage Error (MAPE), and Explained Variation (R^2), skepticism about their broader applicability and reliability persists. A key concern lies in their limited capacity to incorporate domain-specific knowledge, which is essential to ensure interpretability and trustworthiness, particularly in critical fields such as healthcare, environmental science, and finance. These models are often criticized as "black-box" approaches due to their inherent opacity, making it challenging to discern how input variables influence outputs. Also, LSTMs require careful tuning of parameters like learning rate and dropout rate, which can be time-consuming. To address these limitations, recent efforts have focused on advancing explainable AI techniques and developing methods to interpret neural network models [103, 104, 105, 106, 107].

When the focus is narrowed to LSTM models, interpretability becomes even more crucial, given their widespread use in time series forecasting and sequence modeling. Methods aimed at improving the transparency of LSTM models can be broadly categorized into two main approaches. The first involves intrinsically explainable models that integrate interpretability directly within the network architecture. The second relies on *post hoc* explanation techniques, which extract insights from pre-trained models using methods such as structure-based propagation or surrogate local models (e.g., linear approximations).

In the first category, the temporal nature of LSTM algorithms has inspired the

adoption of self-attention mechanisms to improve interpretability. For example, [108] proposed a LSTM model augmented with an attention mechanism to derive information in variables for explanatory purposes. Similarly, [109] incorporated a self-attention module into a framework LSTM to improve interpretability in mixed input time series forecasting. Furthermore, [110] introduced an interpretable memristive[1] LSTM model, demonstrating that the insights extracted from this approach align closely with domain expert knowledge in the built environment.

In the second category, *post hoc* techniques have been utilized to make LSTM models more transparent. For example, [112] calculated the impact value of features by integrating gradients from a LSTM model in a hydrological application. Similarly, [113] and [114] used layer-wise propagation methods to improve interpretability. For an LSTM model designed for the classification of DDoS attacks, [115] utilized LIME, SHAP, Anchor, and LORE to extract the importance of features, using input perturbations to analyze the model predictions. Additionally, in a multi-step ARIMA–LSTM model for tuberculosis incidence forecasting [116], demonstrated that SHAP-based feature importance[2] aligns with conclusions drawn from domain studies. Together, these approaches represent significant advancements in making LSTM models more transparent and interpretable, thereby addressing the broader challenge of enhancing trust and usability in deep learning applications.

GAMs for environmental health risk assessment provide association coefficients

[1] A *memristive model* is a mathematical representation of a memristor, which is a theoretical electrical component where the resistance depends on the history of the electric charge that has flowed through it, essentially acting like a "memory resistor" - meaning its resistance changes based on past current flow, allowing it to store information [111, 110].

[2] SHAP being the implementation of Shapley values, an idea from economics and game theory [117] that has made its way into machine learning as a tool for interpretability.

between log mortality and air pollutant levels, offering valuable insights for policy-making – such as determining thresholds for pollutant concentrations to ensure public safety. By contrast, the proposed LSTM model from Chapter 3, while effective, lacks interpretability for credibility verification and does not produce comparable coefficients for reference. To bridge this gap, this study applies the LIME method to the existing LSTM model, utilizing surrogate local models to interpret feature importance and assess their impact on mortality.

4.2 LIME method for tabular LSTM regression

Regression models establish quantitative relationships between predictors (independent variables) and responses (dependent variables). Statistical regression models, such as linear regression, logistic regression and the general structure of Generalized Linear Models (GLMs) and Generalized Additive Models (GAMs), rely on clear assumptions about these relationships. The coefficients then derived from such models offer interpretable insights for further investigation.

For example, one of the important tasks for GAM models in general air pollution epidemiology studies is to analyze the relationship between air pollution exposure and health outcomes. Quantitative evaluation of the adverse impacts from air pollution on public health can facilitate future regulation of air pollutant emission, environmental policymaking, development of social healthcare, etc. In a standard GAM used for short-term air pollution-related health risk assessment (Eqn. (3.14) in Chapter 3), its main component is the air pollutant predictor. Natural cubic

regression spline functions are used as smoothers for temperature and time to accommodate the measured covariate and unmeasured confounders such as seasonal variation and other time-dependent predictors that may influence the health outcome of interest. This health risk assessment model is rewritten in Eqn. (4.1) to highlight the different components and the exposure-response relation.

$$\log(\mu_t) = \beta_0 + \beta_1 \chi_t + \beta_2 \text{DOW}_t + \underbrace{\text{NS}(T, df)_t + \text{NS}(\text{Time}, df)_t}_{\text{Smoothers}} \tag{4.1}$$

In Eqn. (4.1), β_1 is the parameter of interest, reflecting the air pollution exposure and health response association under the assumption of generalized linearity. If the assumptions are met, it should measure the rate at which the logarithm of health outcome μ_t (such as daily mortality) changes per unit increase of concentration level of air pollutant χ_t.

By contrast, deep learning models like the LSTM model proposed in the previous chapter function as black boxes, producing outputs for any given input once trained, without inherently providing explanations for the associations between inputs and outputs. When a deeper understanding of these relationships is required – particularly for a specific localized input – a surrogate locally interpretable model, such as linear or logistic regression, can be constructed. This approach involves using the localized input, along with perturbations in its vicinity, and their corresponding responses predicted by the black-box model. The surrogate model is then fitted to this data, offering local explanations for the black-box model's behavior.

This concept forms the basis of the LIME method. The term "vicinity" typically refers to a defined measure of distance. For tabular input data, as in the case of

our environmental health risk assessment LSTM model, this distance can be easily defined, often using the Euclidean distance.

In addition, the LIME method has the potential to handle temporal dependencies, as each date in the dataset can be seen as an instance.

The steps involved in LIME are:

- **Select an Instance**: Choose the specific data point you want to explain. For example, in our LSTM model with air pollutant PM_{10}, O_3 and CO to predict daily mortality, we select a specific date, Jan 1st, 2000, and the inputs of that date as an instance.

- **Create Perturbations**: Generate a set of new instances by slightly altering the features of the selected instance. This step creates a neighbourhood of similar data points. For example, if the selected instance includes features like age, income, and tenure, you might create perturbed instances by adding or subtracting small random values to these features.

- **Make Predictions**: Use the original complex model to make predictions on these perturbed instances. This step helps to understand how the model behaves in the neighbourhood of the selected instance.

- **Weight the Perturbed Instances**: Assign weights to these perturbed instances based on their similarity to the original instance. Typically, a proximity measure (e.g., Euclidean distance) is used, where closer instances receive higher weights. This step ensures that the explanation is focused on the local area

around the selected instance.

- **Train an Interpretable Model**: Fit a simple, interpretable model (like a linear regression or decision tree) to the weighted perturbed instances. This model approximates the behavior of the complex model locally. The simplicity of this model allows it to be easily interpreted by humans.

4.3 Data Preparation

Similarly to the Chicago dataset that has been used in previous chapters, data of 10 populous cities are selected from the 108 U.S. cities in the NMMAPS database, and 9 Census divisions (CDS) in Canada are also prepared for the following experiments and demonstration of the utility of LIME.

The 10 cities in the NMMAPS database are Los Angeles, New York, Dallas/Fort Worth, Houston, San Diego, Miami, San Bernardino, San José, Riverside, and Philadelphia. The reasons for selecting these cities and details in the preparation of the data are as follows.

1) The population of each selected city was more than 1.3 million by the year 2000, which can facilitate the interpretable LSTM model analysis of the air pollution-related health risk.
2) Considering that not all cities in the NMMAPS database have complete measurement of the air pollutants of interest, these cities with relatively less missing data are selected for analysis.
3) All-cause non-accidental daily mortality is evaluated with PM_{10} and O_3 as the

air pollutants of interest. For each city, the daily non-accidental mortality data for three age categories (under 65, 65 to 74, and above 75) are aggregated as the daily mortality observation, similar to the preparation of Chicago dataset.

4) Missing values of the air pollutants and the health outcome for each city are interpolated using a second-order polynomial function, and the outliers are dealt with as in experiments of previous chapters.

5) Most PM_{10} concentration levels in the NMMAPS database were recorded every six days while the O_3 data were collected on a daily basis. Taking into account the formulations of the LSTM model, all these data are reorganized on a 6-day interval, such that the original 14 years data are regrouped into a time series of about 852 periods.

Data from nine Census Divisions (CD) in Canada were collected from public climate and air pollutant datasets, as well as daily mortality data from Health Canada. The CDs are Ottawa (CD3506), Toronto (CD3520), Halton (CD3524), Hamilton (CD3525), Simcoe (CD3543), Alberta No.6 (CD4806; Calgary), Alberta No.11 (CD4811; Edmonton), Greater Vancouver (CD5915), and Capital, BC (CD5917). These CDs were selected on the basis of the availability of consecutive daily measurements for a reasonable period of time. Ottawa (CD code 3506) has data spanning from November 9th, 2006 to December 31st, 2015, with a total of 3340 entries. Toronto (CD code 3520) has data spanning from August 17th, 2003 to December 31st, 2015, with a total of 4520 entries. Halton (CD code 3524) has data spanning from June 24th, 2007 to December 31st, 2015, with a total of 3113 entries. Hamil-

ton (CD code 3525) has data spanning from October 1st, 2005 to December 31st, 2015, with a total of 3744 entries. Simcoe (CD code 3543) has data spanning from December 2nd, 2001 to November 27th, 2013, with a total of 4379 entries. Alberta No.6 (CD code 4806) has data spanning from January 2nd, 1996 to December 31th, 2015, with a total of 7304 entries. Alberta No.11 (CD code 4811) has data spanning from January 2nd, 1996 to December 31th, 2006, with a total of 4016 entries. Greater Vancouver (CD code 5915) has data spanning from January 2nd, 1996 to December 31th, 2015, with a total of 7304 entries. Capital, BC (CD code 5917) has data spanning from August 27th, 2004 to December 31th, 2015, with a total of 4144 entries. All these datasets contain consecutive daily mortality data for the specified time periods. Additionally, 24-hour mean temperatures and air pollutants (include PM_{10} and O_3) are also available (See Table 4.1).

Census Division	**CD Code**	**Date Range**	**Entries**
Ottawa	3506	2006-11-09 to 2015-12-31	3340
Toronto	3520	2003-08-17 to 2015-12-31	4520
Halton	3524	2007-06-24 to 2015-12-31	3113
Hamilton	3525	2005-10-01 to 2015-12-31	3744
Simcoe	3543	2001-12-02 to 2013-11-27	4379
Calgary	4806	1996-01-02 to 2015-12-31	7304
Edmonton	4811	1996-01-02 to 2006-12-31	4016
Greater Vancouver	5915	1996-01-02 to 2015-12-31	7304
Capital, BC	5917	2004-08-27 to 2015-12-31	4144

Table 4.1: Data Summary for 9 Census Divisions (CDs) in Canada

4.4 Interpretable LSTM model analysis with LIME

The LIME method is applied to the LSTM model proposed in Section 3.3.2. Each input of the training data is treated as an instance for calculating LIME values. Therefore, for any LSTM model with m inputs, such as the lag 1 to $m-1$ days and the current day, the LIME values can be calculated for each instance.

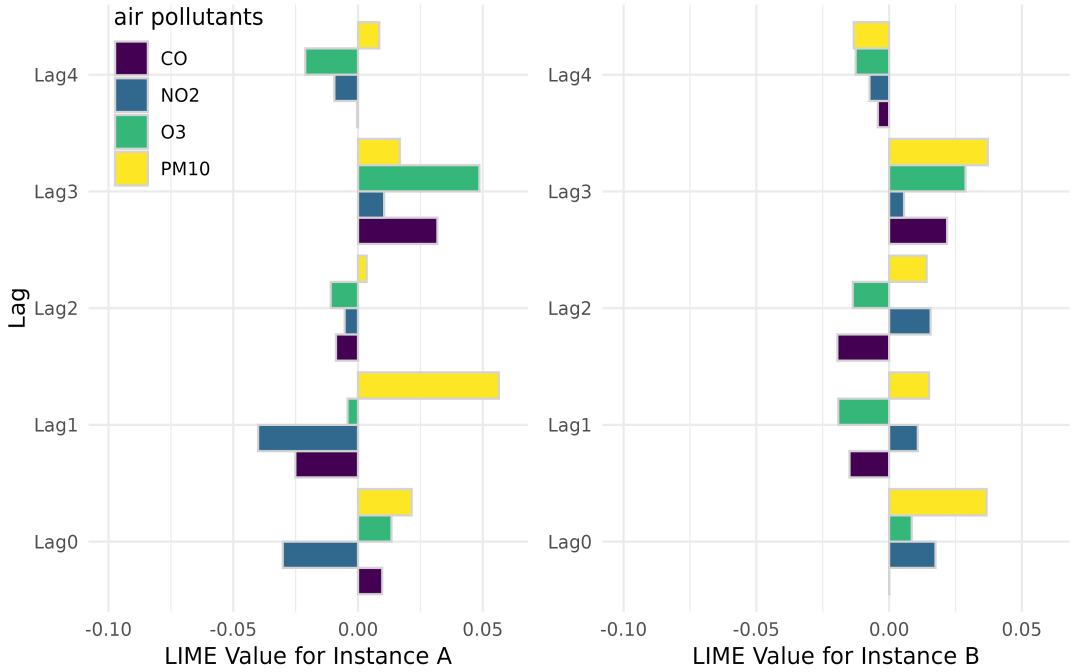

Figure 4.1: The LIME explanation at two local instances in the LSTM model with the Chicago dataset.

For the Chicago LSTM model with $m = 5$ and air pollutants CO, NO_2, O_3 and PM_{10} as features, the LIME values for two instances are showcased in Fig. 4.1. Instance A represents the input data in January 5th, 1987. Instance B represents the input data in December 1st, 1996. In Instance A, PM_{10} has a significantly positive LIME value for Lag 1. In Instance B, PM_{10} also shows large importances for Lag 0 and Lag 4 (illustrated by the purple bars). O_3 displays a large importance for Lag 3

in both Instance A and Instance B (depicted by the green-blue bars). LIME values and feature importances are also calculated for other training data. Figures 4.2, 4.3, 4.4, 4.5, and 4.6 present the LIME values for all instances in the Chicago LSTM model. The LIME values for each feature or air pollutant are plotted as lines over time.

In Figure 4.7, we present the *monthly* LIME values of the Chicago LSTM model for Lag 0. The LIME values for air pollutants are illustrated in box plots, with their monthly averages represented by yellow lines. Figure 4.8 shows the yearly average LIME values for O_3 from Lag 0 to Lag 4. The gray bands represent the smoothed range of the LIME values.

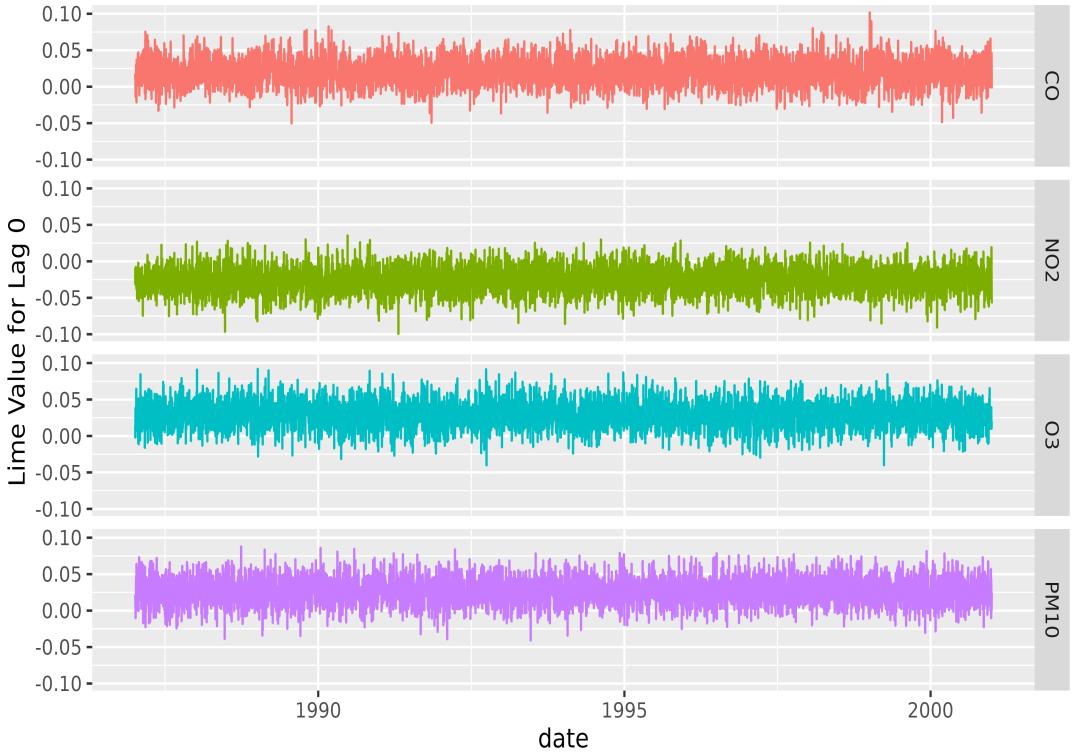

Figure 4.2: The LIME values for 4 air pollutions with Lag 0 in the LSTM model with the Chicago dataset.

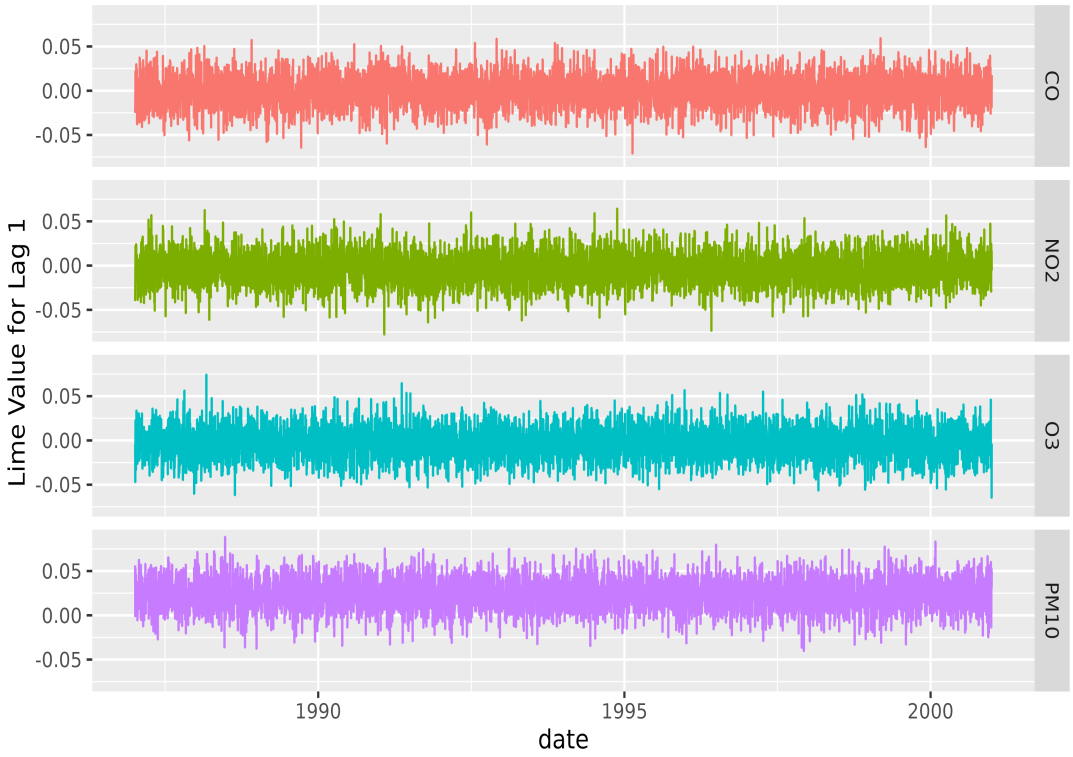

Figure 4.3: The LIME values for 4 air pollutions with Lag 1 in the LSTM model with the Chicago dataset.

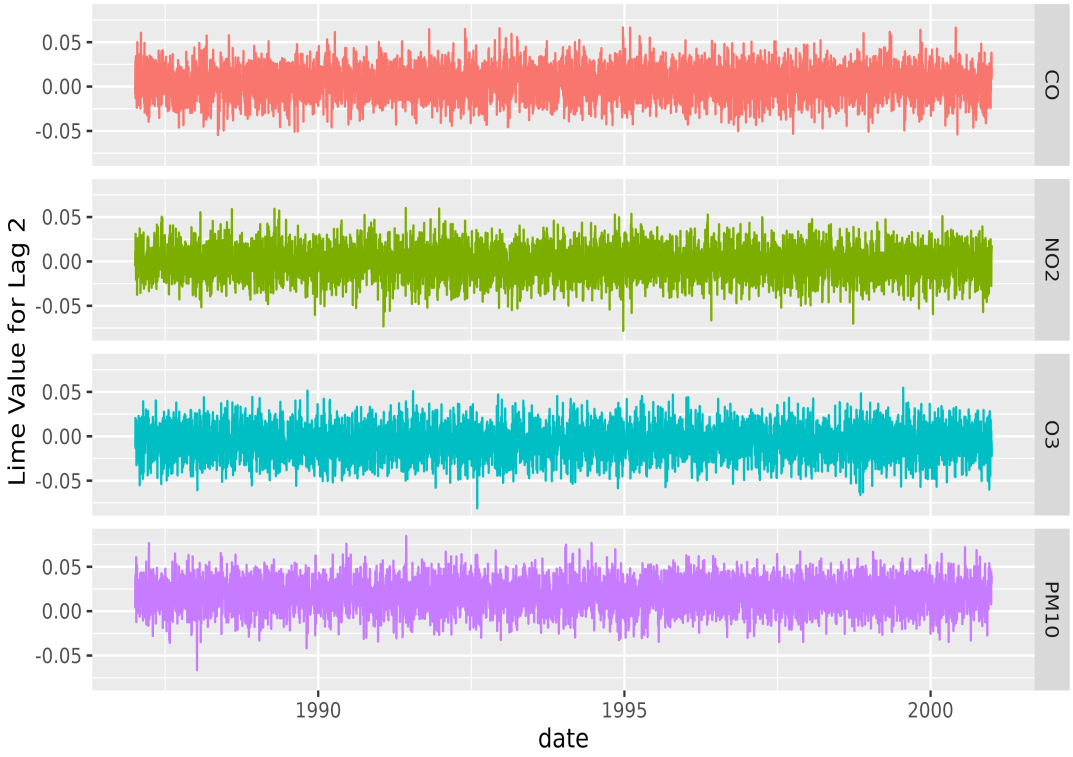

Figure 4.4: The LIME values for 4 air pollutions with Lag 2 in the LSTM model with the Chicago dataset.

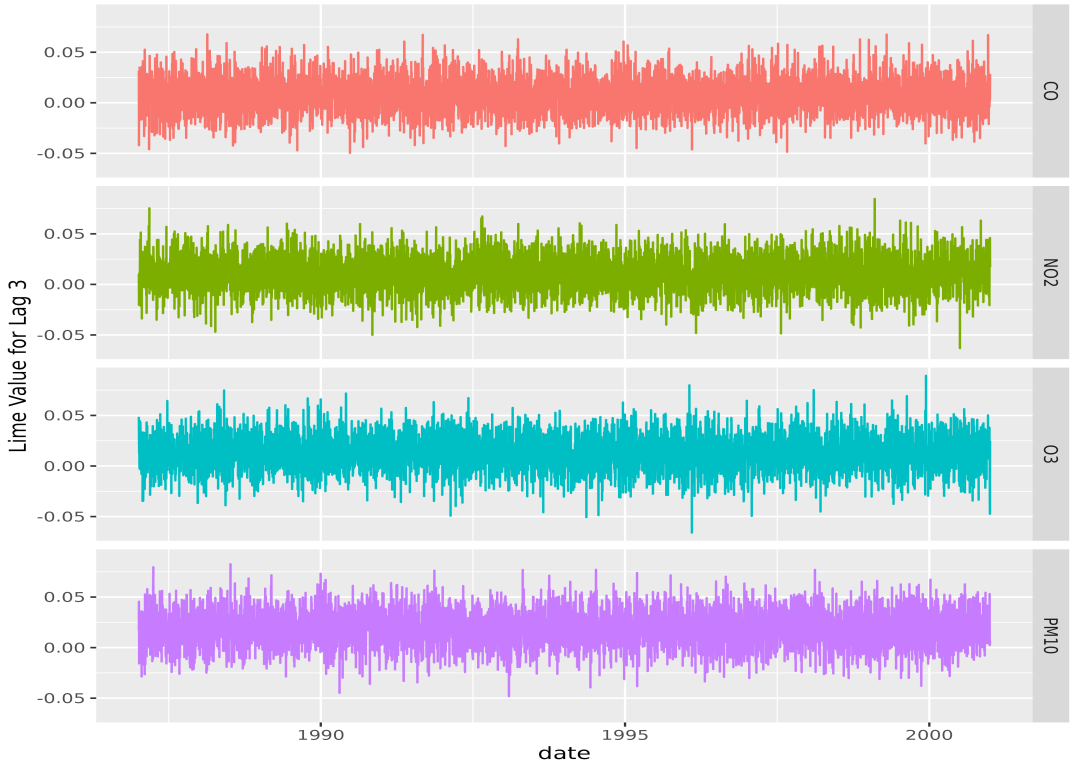

Figure 4.5: The LIME values for 4 air pollutions with Lag 3 in the LSTM model with the Chicago dataset.

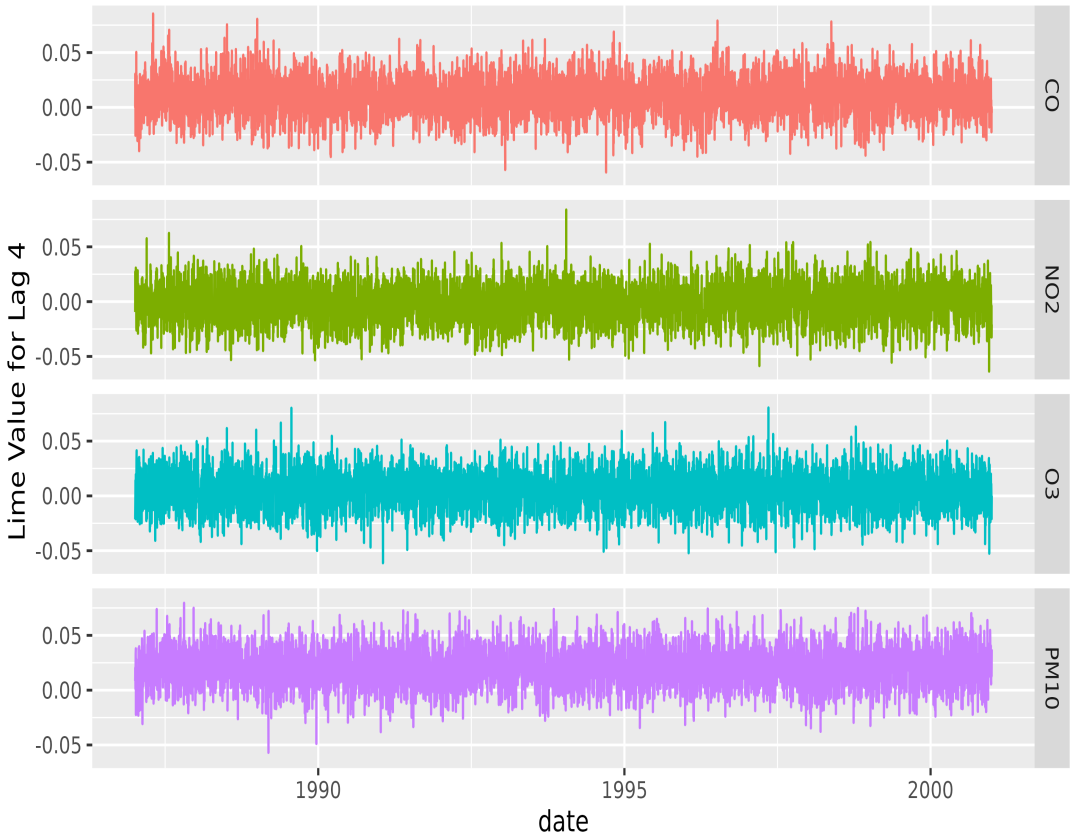

Figure 4.6: The LIME values for 4 air pollutions with Lag 4 in the LSTM model with the Chicago dataset.

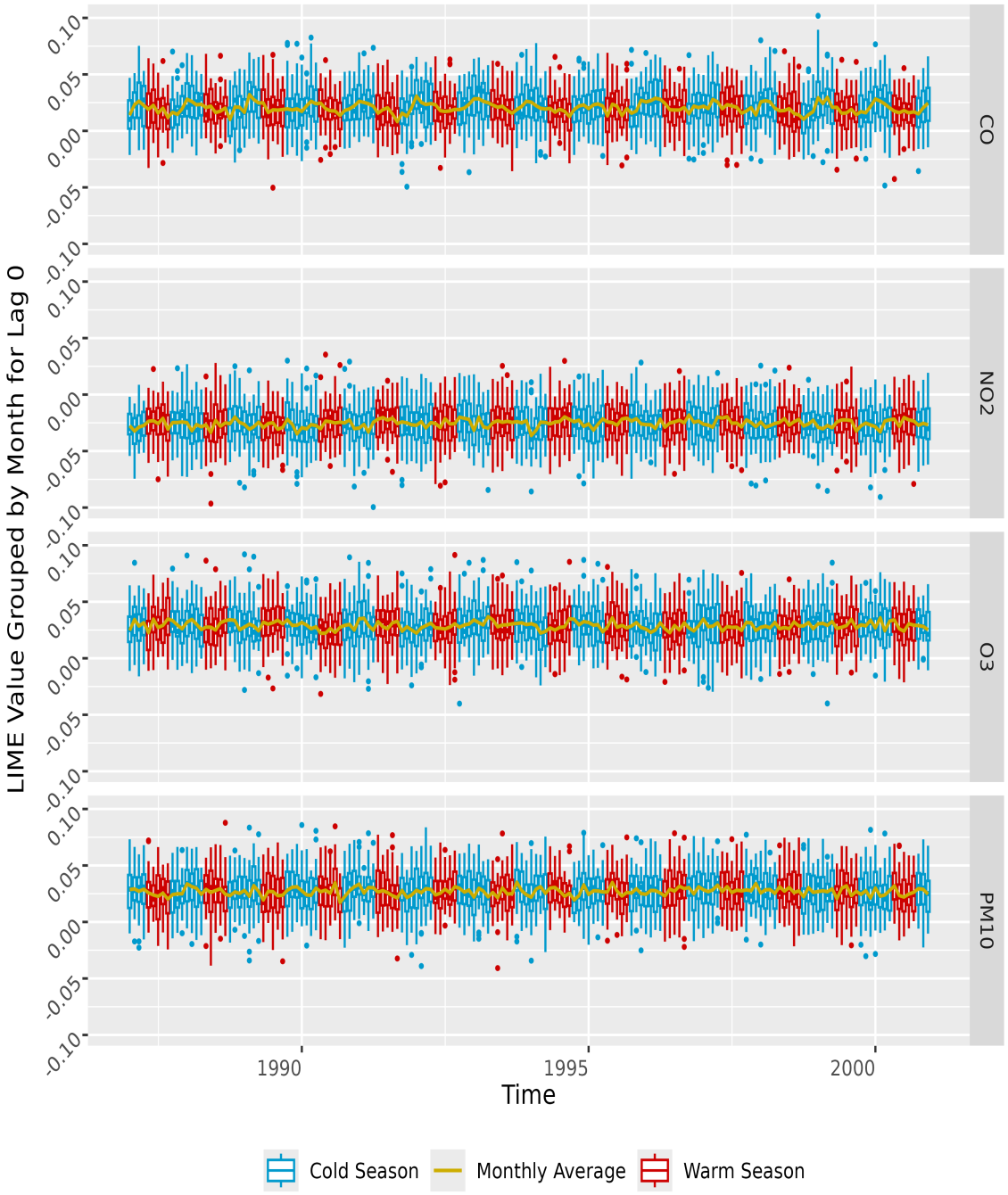

Figure 4.7: The monthly LIME values for 4 air pollutions with Lag 0 in the LSTM model with the Chicago dataset.

Based on the LIME values for the Chicago LSTM model, the relative importance of lag values was analyzed. For the feature CO, lag 0 and lag 3 were found to be more important compared to lag 1, lag 2, and lag 4. Similarly, for features NO_2, O_3,

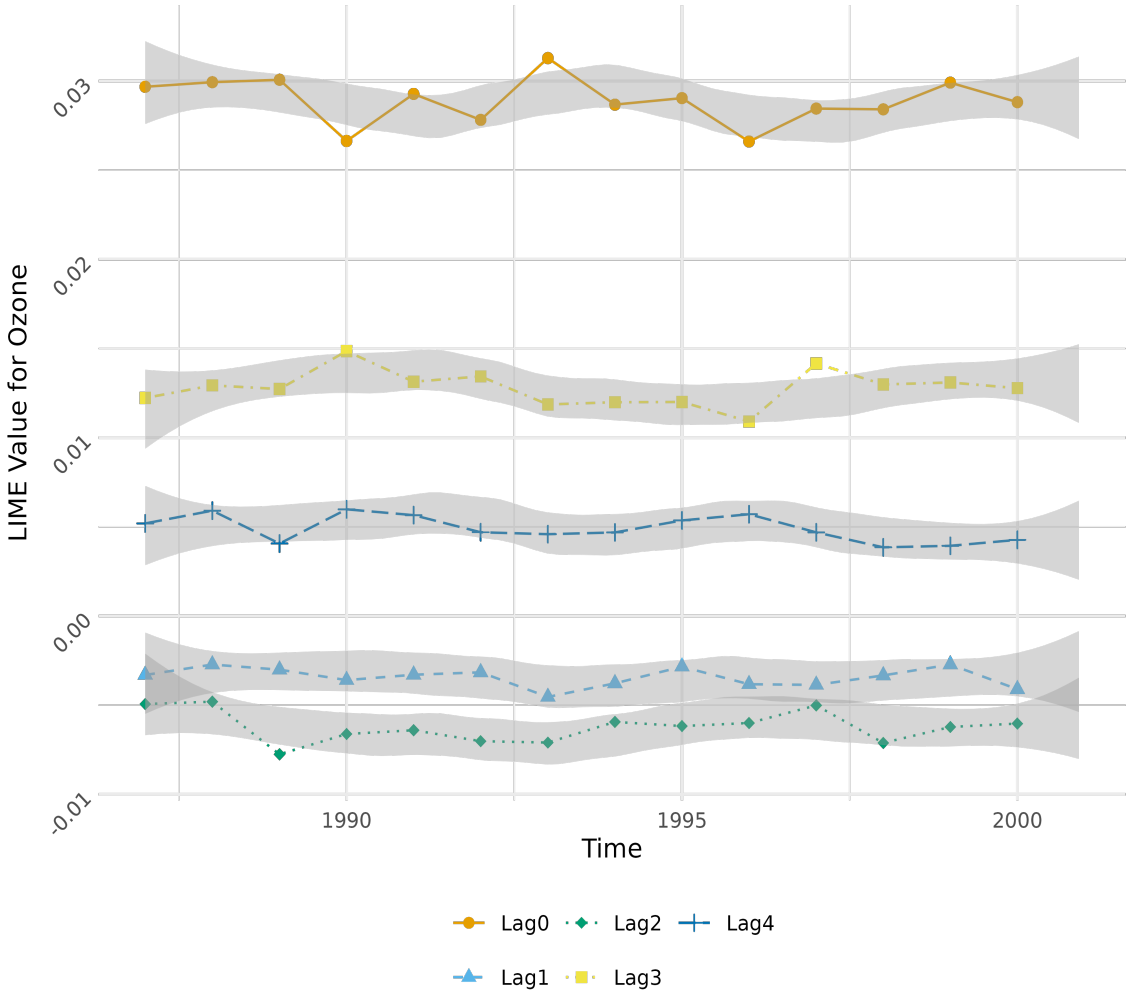

Figure 4.8: The Yearly Average LIME values for O3 in the LSTM model with the Chicago dataset.

Air Pollutant	Lag0	Lag1	Lag2	Lag3	Lag4
CO	0.0204672	-0.0004625	0.0043803	0.0078123	0.0099915
NO2	-0.0255944	-0.0038639	-0.0013523	0.0105443	-0.0014620
O3	0.0289017	-0.0034358	-0.0062252	0.0127930	0.0048987
PM10	0.0269475	0.0236706	0.0183971	0.0188985	0.0203165

Table 4.2: Average LIME values in the Chicago LSTM model with lag 0 to lag 4

and PM_{10}, lag 0 and lag 3 demonstrated greater significance than lag 1, lag 2, and lag 4 (refer to Table 4.2). Consequently, the LSTM model was refitted using only lag 0 and lag 3 as input variables. The average LIME values for the new model were

Air Pollutant	Lag0	Lag3
CO	0.0384596	-0.0526886
NO2	-0.0226117	-0.0204034
O3	0.0243523	-0.0122284
PM10	0.0363861	0.0451869

Table 4.3: Average LIME values in Chicago LSTM model with lag 0 and lag 3

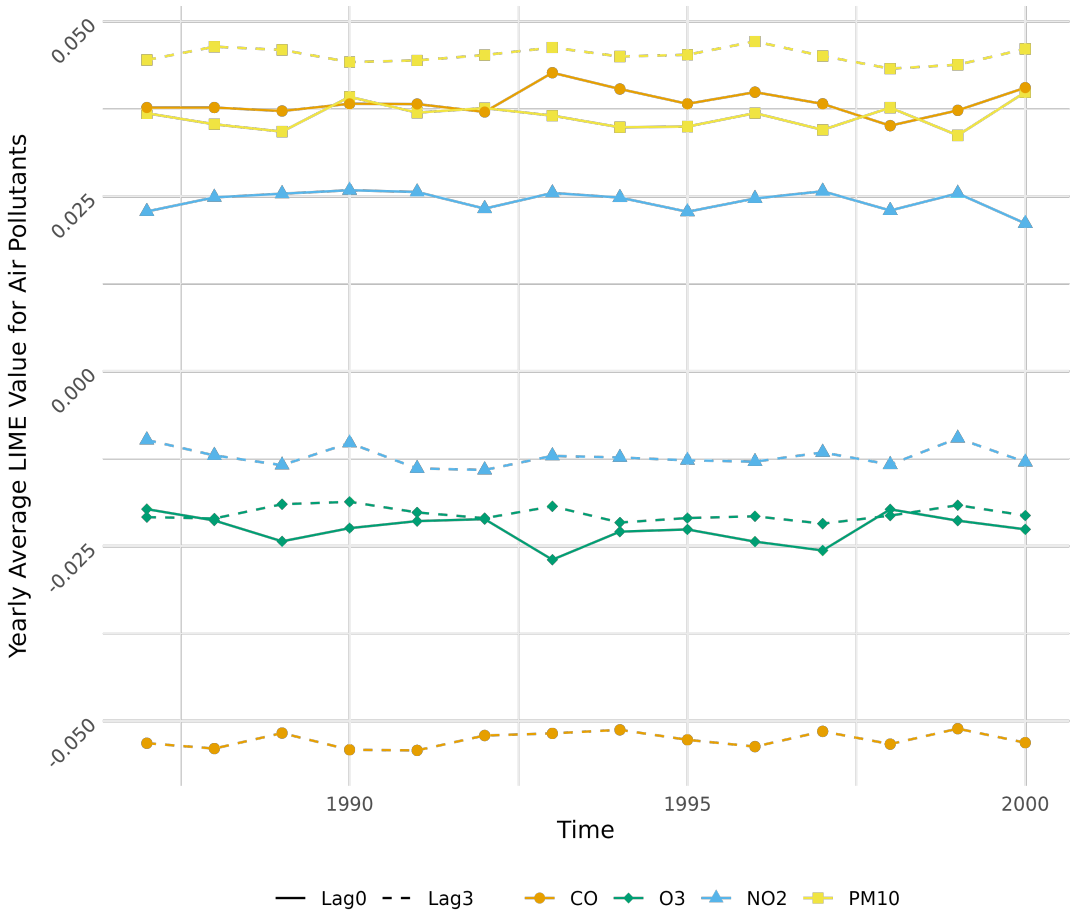

Figure 4.9: The Yearly Average LIME values for the LSTM model with only Lag 0 and Lag 3.

consistent with those of the original model in terms of scale (see Table 4.3). Figures 4.9 illustrate the yearly average LIME values of the new model.

To further streamline the analysis, LIME values for lag 0 and lag 3 were combined, producing a single yearly average LIME value to represent the air pollutant's

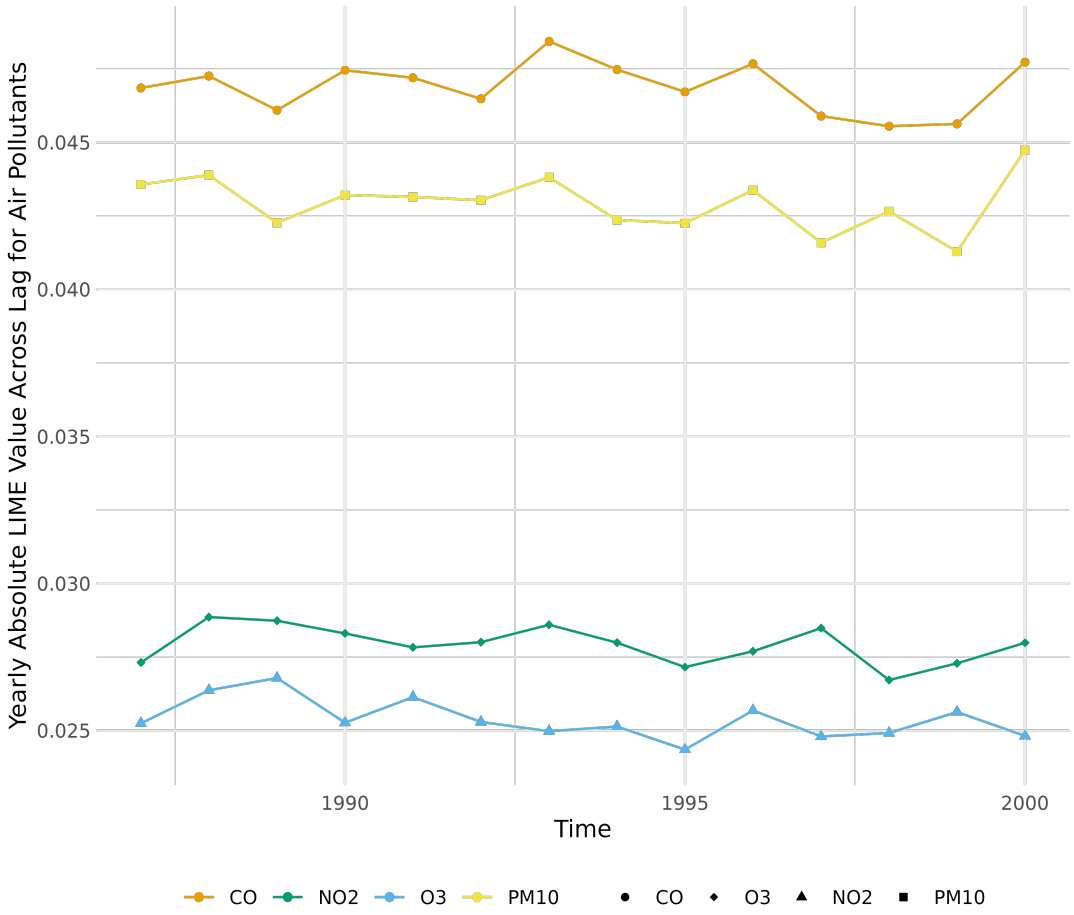

Figure 4.10: The Yearly Absolute Average LIME values Across Lag 0 and Lag 3.

importance in the model. Given that LIME values include both positive and negative numbers, using a simple average could result in positive and negative contributions offsetting each other. Therefore, the absolute LIME values from lag 0 and lag 3 were averaged to create a synthetic importance metric. The results of this approach are displayed in Figure 4.10.

4.4.1 Sensitivity on Lag for LSTM models

To evaluate the sensitivity of the model training process to the inherently stochastic nature of the solvers used, we aggregated the results from multiple independent

training iterations. This method allows us to account for the stochastic nature of the training process, including variations introduced by random initialization, data shuffling, and optimization dynamics. The analysis focused on the LSTM model introduced in Chapter 3, applied to the Chicago dataset derived from NMMAPS.

We computed the density distributions of key performance metrics to provide a robust depiction of model behavior. Specifically: RMSE, MSE, MAE, MAPE, and R^2 were examined, and their distributions are presented in Figures 4.11, 4.12, 4.13, 4.14, and 4.15, respectively.

The results reveal a consistent shift toward zero for the modes of the density distributions of these metrics as the input lags included increase (that is, as m increases). This trend indicates a reduction in error and an improvement in predictive accuracy, affirming the hypothesis that incorporating more lagged information into the LSTM models enhances their ability to capture temporal dependencies in the data. The findings corroborate the theoretical advantages of increasing input length in sequence-based models: doing so reduces the error in the predictive fit of the models. Which, on some level, matches our general understanding of time series models where including more lags almost always gives a better "prediction", regardless of whether there are true covariances present.

Figures 4.11, 4.12, 4.13 and 4.14 all show variations of the measurement of error, and consistently show decreases of the mode with increasing m. However, Figure 4.15 shows the R^2, i.e., the explained variation, and it shows increasing modes. These results match nicely, and indicate that for increasing predictive power, more lags con-

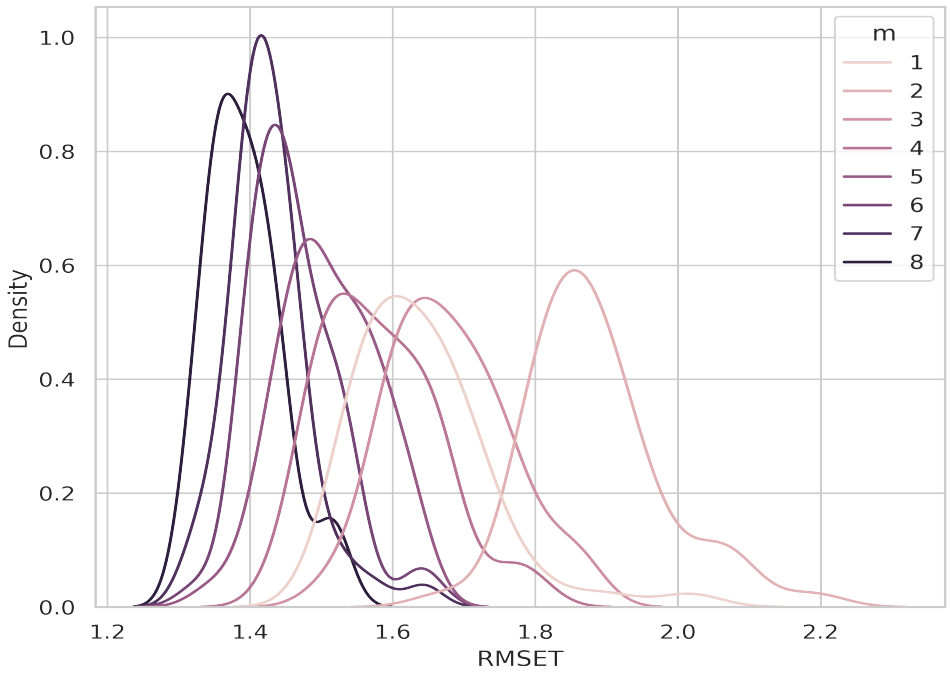

Figure 4.11: The RMSE density on test set for 50 times of training on a LSTM model.

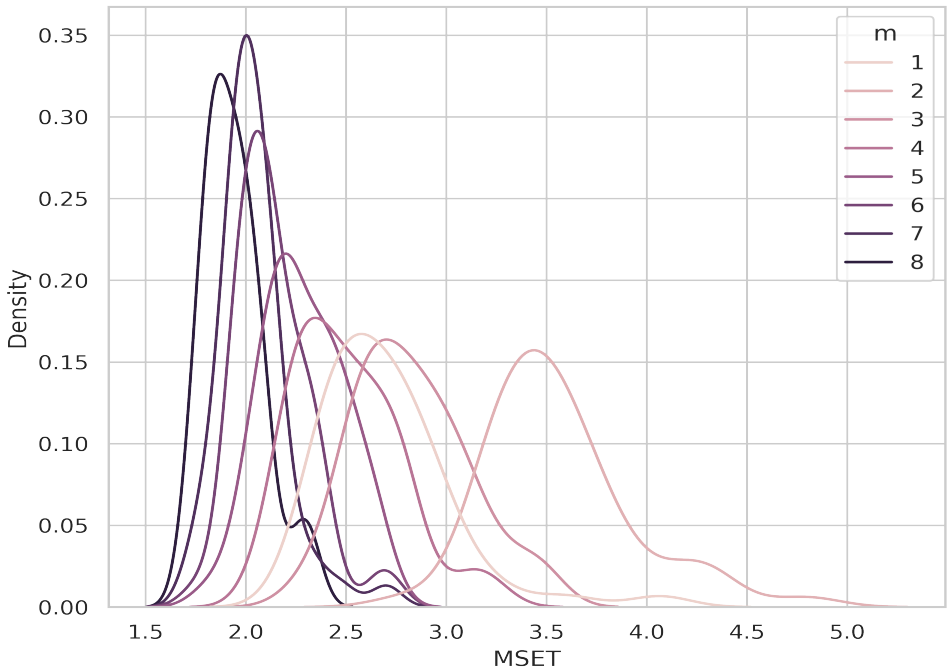

Figure 4.12: The MSE density on test set for 50 times of training on a LSTM model.

tribute to success. There is some evidence in Figures 4.11-4.13 that the increasing lags begin to have less and less influence after $m = 6$, which also matches our pre-

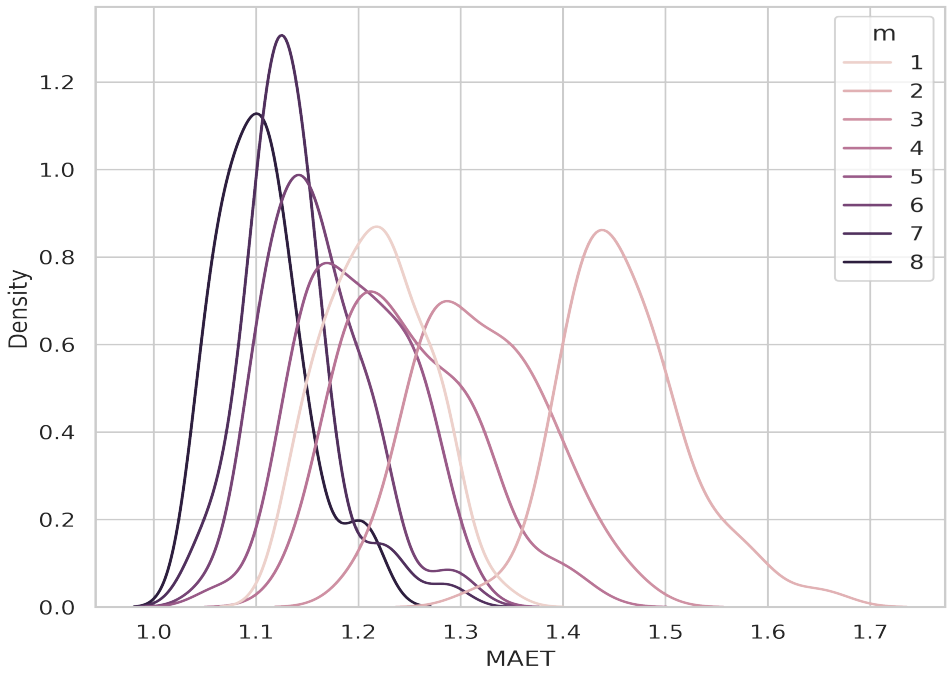

Figure 4.13: The MAE density on test set for 50 times of training on a LSTM model.

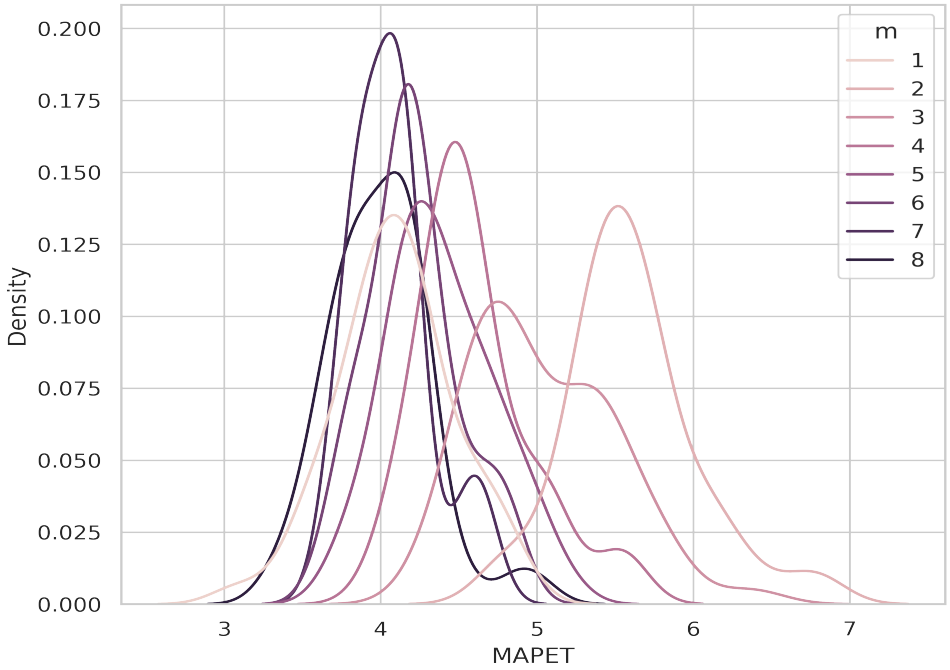

Figure 4.14: The MAPE density on test set for 50 times of training on a LSTM model.

viously understood knowledge about time series-based models and the explanatory power of covariance structures.

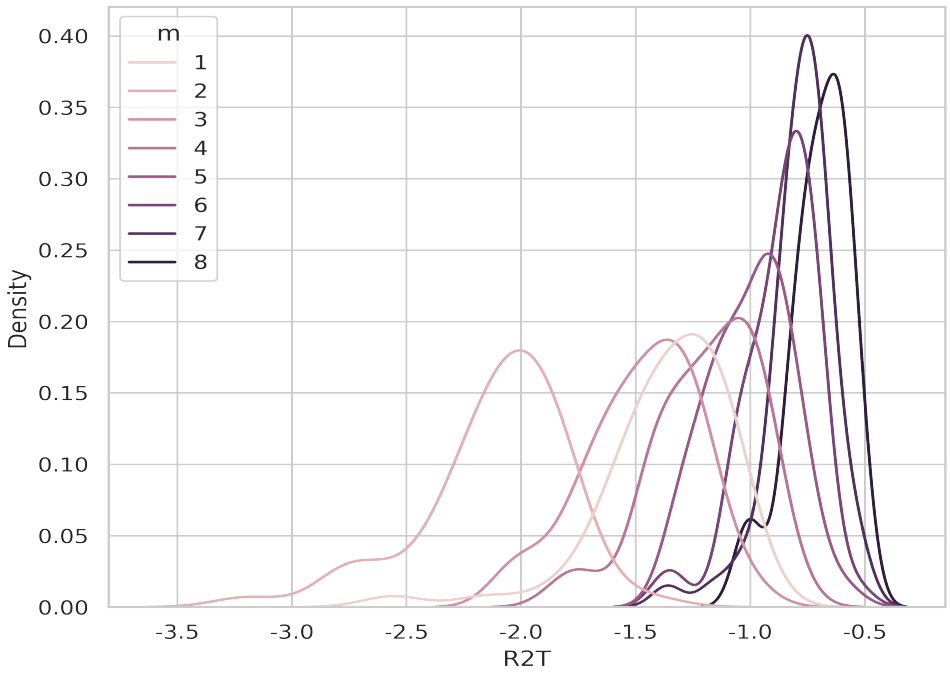

Figure 4.15: The R2 density on test set for 50 times of training on a LSTM model.

4.4.2 Analysis for Lag Significance

One of the key advantages of LSTM modeling in explaining health-environment association issues is its ability to account for distributed lags within the model structure. Unlike the widely used GAM models, which typically require explicit specification of lag terms, LSTM models inherently capture temporal dependencies across various time steps, enabling them to model complex, non-linear relationships between environmental exposures and health outcomes over time. This ability to automatically incorporate distributed lags allows LSTM models to more effectively capture delayed effects, which are often observed in health-environment studies. That is, at least, more effectively in the sense of having higher explanatory power for the response – explained variation being quite high. The challenge, as discussed at the end of Section 3, is in how to interpret these complex nonlinear models and their good

predictive power.

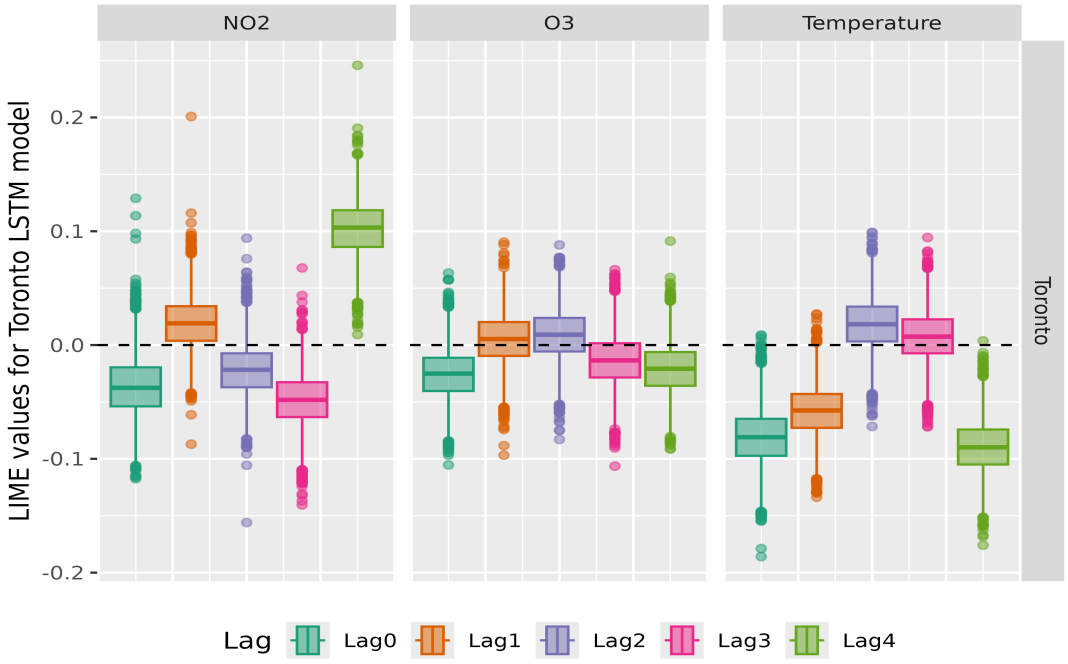

Figure 4.16: The LIME values in LSTM model with m=5 for the Toronto datasets.

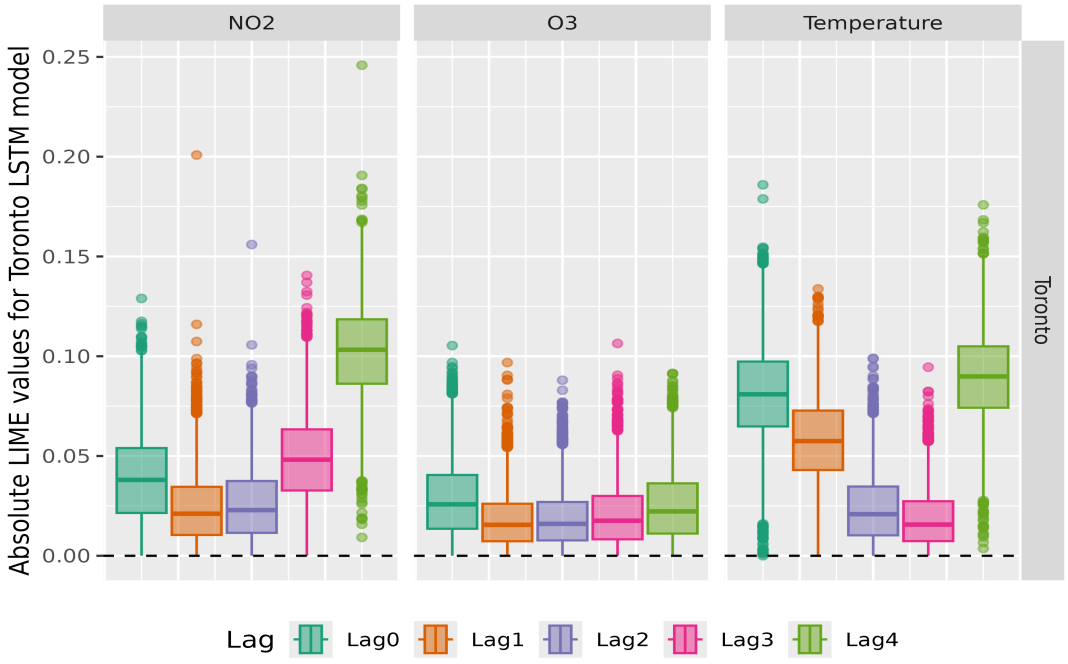

Figure 4.17: The absolute LIME values in LSTM model with m=5 for the Toronto datasets.

Predictor	Lag	Mean	Median	Lower 95% CI	Upper 95% CI
	Lag0	-0.036	-0.038	-0.037	-0.036
	Lag1	0.019	0.019	0.018	0.020
NO2	Lag2	-0.022	-0.022	-0.023	-0.021
	Lag3	-0.048	-0.048	-0.049	-0.048
	Lag4	0.102	0.103	0.102	0.103
	Lag0	-0.026	-0.025	-0.026	-0.025
	Lag1	0.005	0.005	0.005	0.006
O3	Lag2	0.009	0.009	0.009	0.010
	Lag3	-0.014	-0.014	-0.014	-0.013
	Lag4	-0.021	-0.021	-0.021	-0.020
	Lag0	-0.081	-0.081	-0.082	-0.080
	Lag1	-0.058	-0.058	-0.059	-0.057
Temp	Lag2	0.018	0.018	0.018	0.019
	Lag3	0.008	0.007	0.007	0.008
	Lag4	-0.089	-0.090	-0.090	-0.089

Table 4.4: Summary of Statistics for LIME values across Different Lags in Toronto datasets

Predictor	Lag	Mean	Median	Lower 95% CI	Upper 95% CI
	Lag0	0.039	0.038	0.038	0.040
	Lag1	0.024	0.021	0.024	0.025
NO2	Lag2	0.026	0.023	0.025	0.026
	Lag3	0.049	0.048	0.048	0.049
	Lag4	0.102	0.103	0.102	0.103
	Lag0	0.028	0.026	0.028	0.029
	Lag1	0.018	0.016	0.018	0.019
O3	Lag2	0.019	0.016	0.019	0.019
	Lag3	0.021	0.018	0.020	0.021
	Lag4	0.025	0.022	0.025	0.026
	Lag0	0.081	0.081	0.080	0.082
	Lag1	0.058	0.057	0.057	0.059
Temp	Lag2	0.024	0.021	0.023	0.024
	Lag3	0.019	0.016	0.018	0.019
	Lag4	0.089	0.090	0.089	0.090

Table 4.5: Summary of Statistics for absolute LIME values across Different Lags in Toronto datasets

To address the explanatory challenges inherent in multi-lag models, we continue to rely on LIME values, with particular emphasis on those associated with different

lag periods.

In the case of the Toronto dataset, which includes lag values ranging from 0 to 4 and air pollutants NO_2 and O_3, we examined the distribution of LIME values across air pollutants and temperature. As illustrated in Figure 4.16 and Table 4.4, the average LIME values for Lag 0 tend to deviate further from the zero baseline compared to those of Lags 1, 2, and 3 in most instances. Additionally, LIME values for Lag 4 exhibit notable prominence in relation to NO_2 and temperature, whether in a positive or negative direction.

Furthermore, if absolute LIME values are considered as an alternative measure of significance within LSTM models—disregarding whether they indicate positive or negative influence – the results (see Figure 4.17 and Table 4.5) indicate that, overall, Lag 0 holds greater importance than Lags 1 through 3. However, Lag 4 emerges as the most significant for NO_2 and temperature.

Similarly, we run the experiments on other CDs from the Canadian datasets. Figures 4.18 to 4.21 present the aggregated absolute LIME values for each lag in the LSTM models, utilizing air pollutants O_3 and PM_{10}, along with temperatures as input features. The lag importance for a single feature in a model can be distinguished using LIME values. It should be noted that the lag differences are not consistent with the increment of m in the LSTM models. As more lagged data is included in the model, the lag importance changes.

Previous studies investigating the lag importance in air pollutant-health assessment research based on traditional models have shown a decreasing importance trend

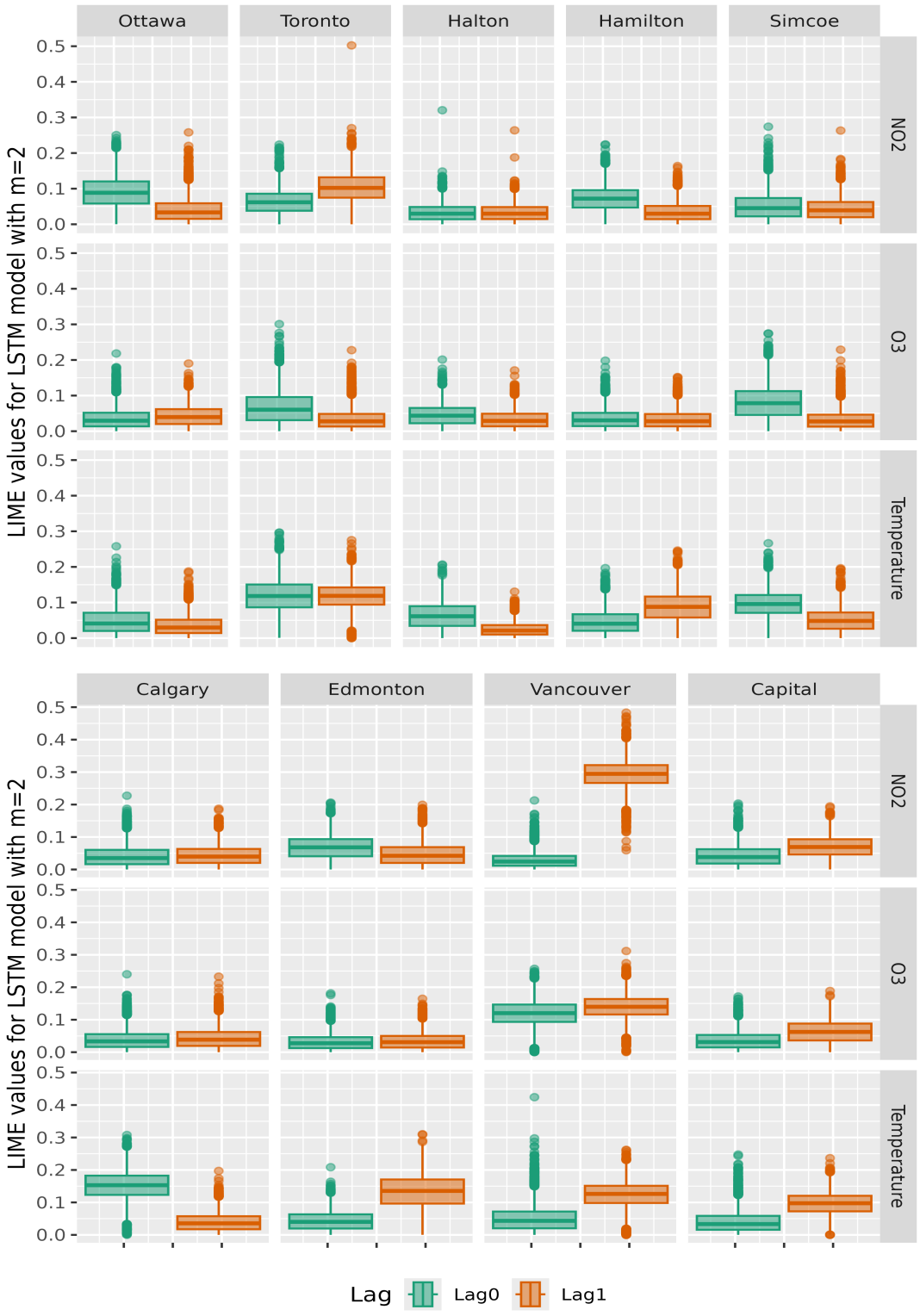

Figure 4.18: The absolute LIME values in LSTM model with m=2 for the Canadian CD datasets.

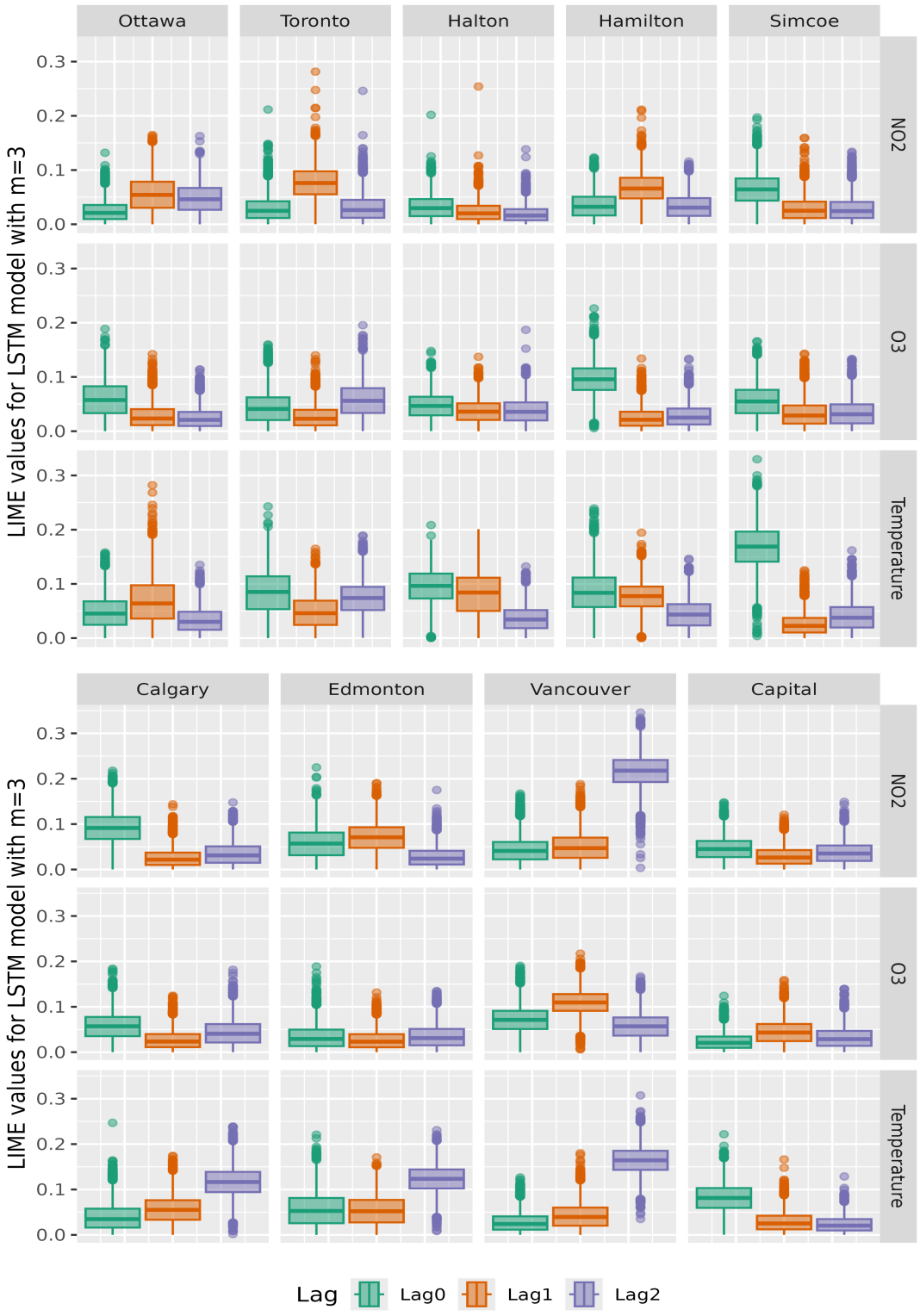

Figure 4.19: The absolute LIME values in LSTM model with m=3 for the Canadian CD datasets.

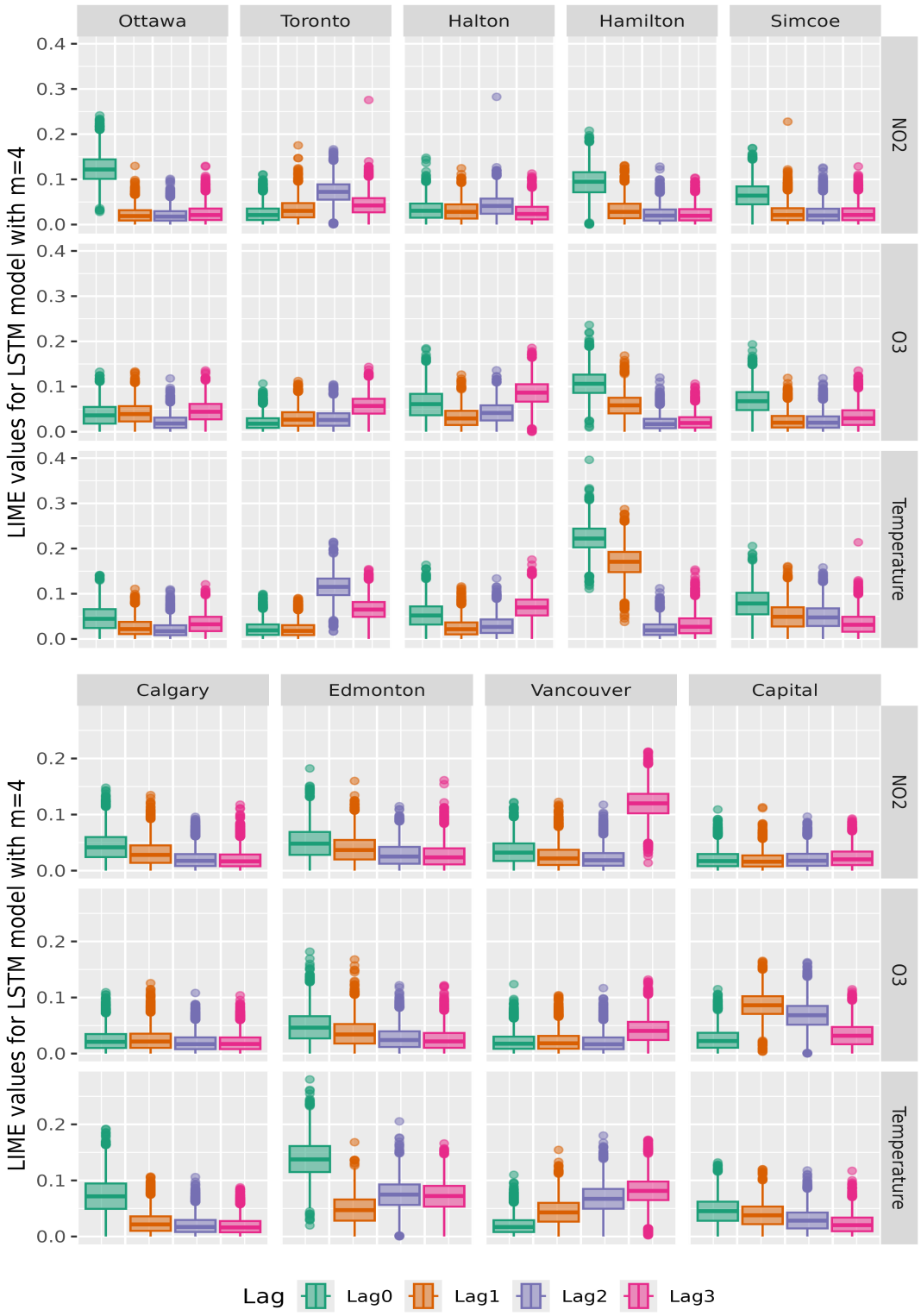

Figure 4.20: The absolute LIME values in LSTM model with m=4 for the Canadian CD datasets.

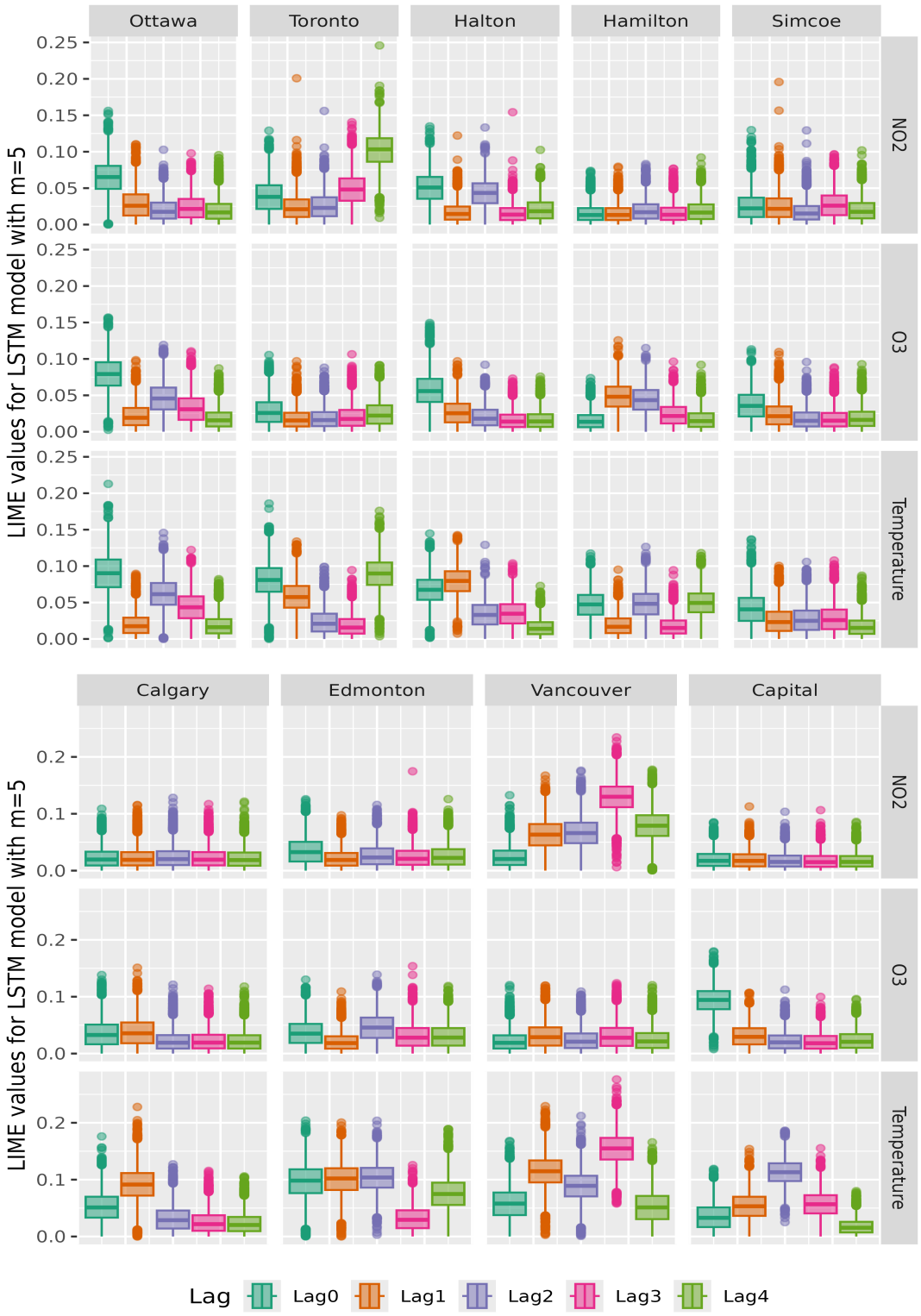

Figure 4.21: The absolute LIME values in LSTM model with m=5 for the Canadian CD datasets.

in the Chicago dataset [70]. Our application of LIME explanations on the LSTM model, using the same data inputs, also reveals a similar decreasing importance trend (demonstrated in Figure 4.22). This indicates that LIME is similarly assigning "importance" to the higher-order lags in line with what previous studies found, i.e., higher lags are less important for explaining the variation in the log health effect response. This is not surprising, but is reassuring that LIME finds the same structures as have been painstakingly extracted from numerous previous studies.

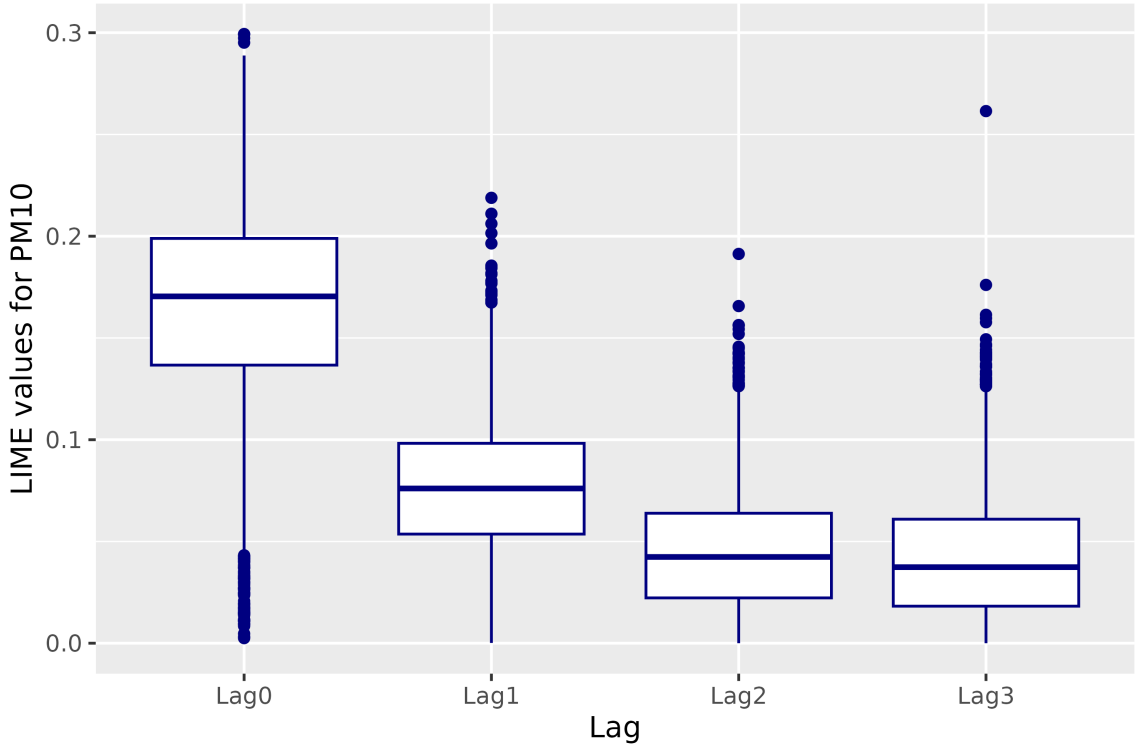

Figure 4.22: The LIME values for PM10 with Lags from 0 to 4 in a LSTM model with the Chicago dataset.

The LSTM+LIME method offers an alternative approach to lag analysis, with results that are consistent with previous studies. Moreover, this novel methodology enables the incorporation of multiple air pollutants within a single model, allowing for lag analysis of each individual input feature. In contrast, prior methods were

limited to considering only one air pollutant at a time.

4.4.3 Analysis for Feature Importance in LSTM model

The LIME values derived from the proposed LSTM model may offer valuable insights into the lagged influence of environmental variables on mortality assessments, as well as the temporal dynamics of feature importance within the model.

For this analysis and demonstration, data from the city of Houston, obtained from the NMMAPS dataset, was utilized. The LSTM model incorporated daily measurements of temperature, PM_{10}, and O_3 as input variables. It should be noted that this dataset contains PM_{10} measurements taken every six days, instead of daily observations. To incorporate this into the tabular regression model, other variables including O_3, temperature, and mortality, are also sampled at six-day intervals. Lagged inputs were considered for the current day and up to four preceding observation days (which are lag 0, lag 6, lag 12, lag 18 and lag 24). The model was trained for 8,000 epochs, after which the LIME values were calculated.

Figure 4.23 illustrates the LIME values corresponding to PM_{10} across various lags, spanning the period from 1987-01-26 to 2000-12-26. The LIME values in warm seasons are in red and the LIME values in cold seasons are represented by blue. The LIME values reveal a clear seasonal pattern: the feature importance of PM_{10} in the LSTM model, as reflected by the LIME values, is generally lower during warm seasons compared to cold seasons. A focused view can be found in Figure 4.24, which depicts PM_{10} LIME values for Lag 0 during the years 1987 and 1988, and

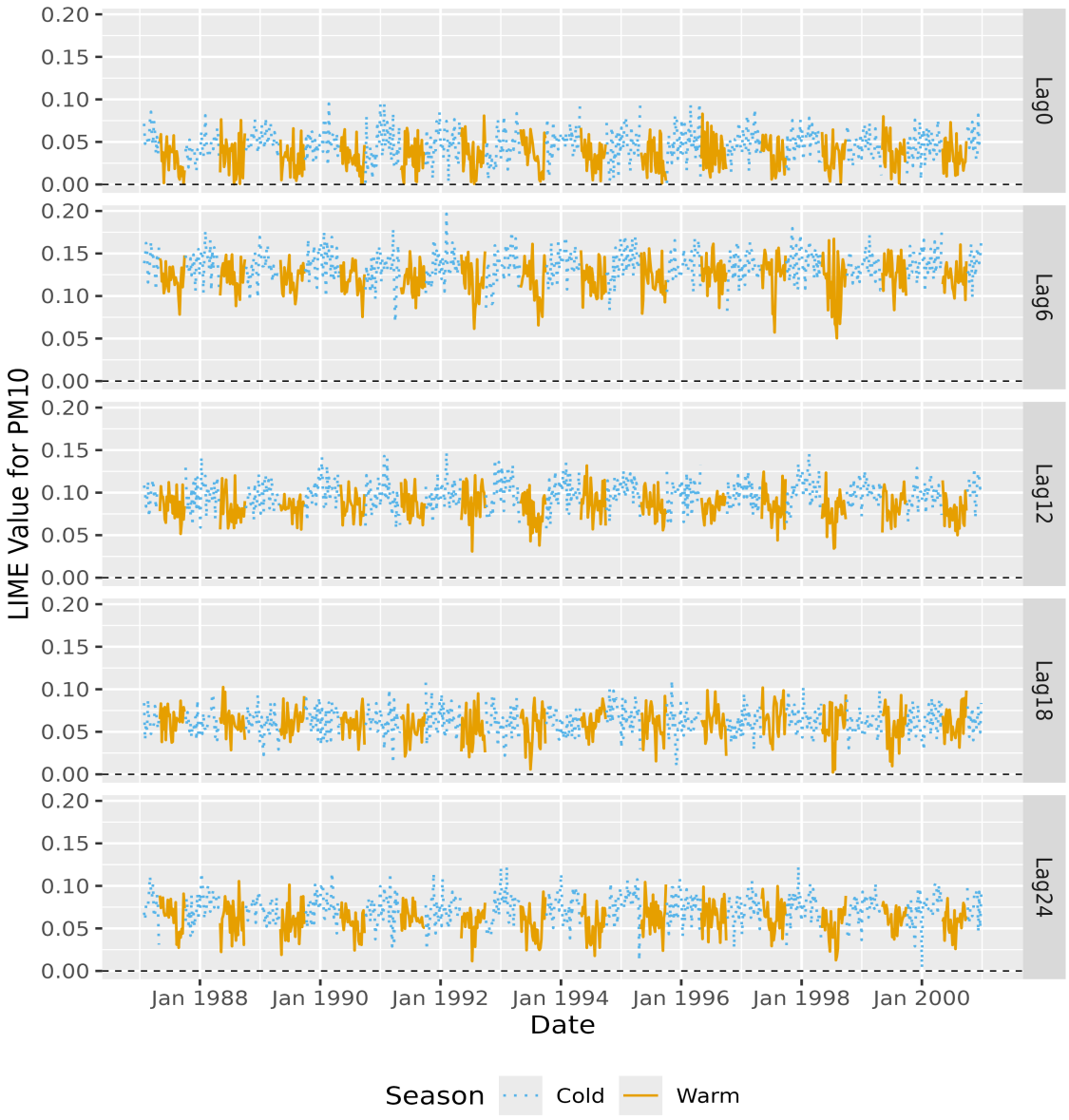

Figure 4.23: The LIME values for PM10 with Lags from 0 to 24 in a LSTM model with Houston dataset.

further highlights this pattern. In warm seasons (red triangles), most LIME values fall within the range of 0 to 0.06, while in cold seasons (blue dots), the majority are in the range of 0.02 to 0.08. The distinction between warm and cold seasons is clearly observed, suggesting a potential relationship between PM_{10} feature importance and daily temperature. When the temperature is warmer, the model considers the PM_{10}

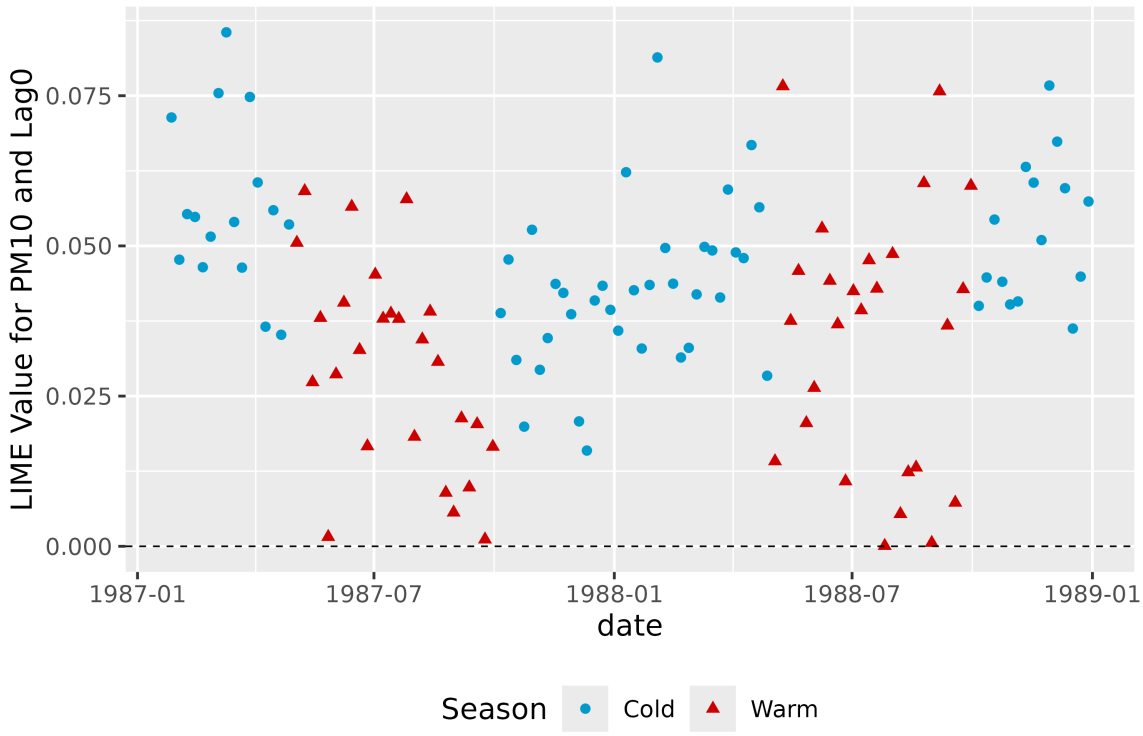

Figure 4.24: The LIME values for PM10 with Lags 0 in a LSTM model applied to the Houston dataset.

to be less important for explaining the variability in daily all-cause mortality; when the temperature is colder, the PM_{10} is more important. This is an interesting finding.

To further explore this relationship, Figure 4.25 presents scatter plots of LIME values for PM10 against daily temperatures in the dataset. Distinct decreasing trends are observed across different lags. The LIME values for Lag 0 are depicted as orange dots, while those for Lag 1 (recall: this is 1 'unit', which means 6 days, so labeled as Lag6 in the plot) are represented by turquoise triangles. Linear models fitted to these two groups of paired data reveal negative slopes (-0.00050 for Lag 0 and -0.00057 for Lag 6), indicating that as temperature increases, the feature importance of PM_{10} in predicting daily mortality using the LSTM model decreases. This finding suggests that on colder days, PM_{10} plays a more significant role in assessing adverse

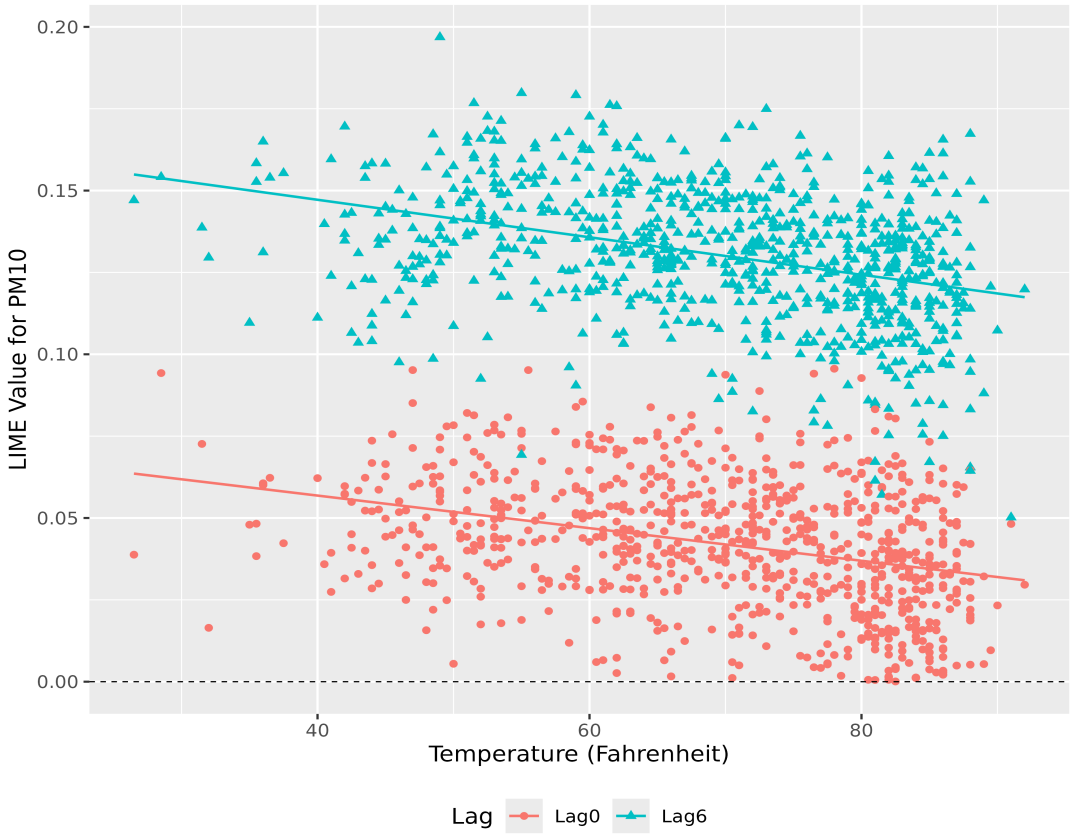

Figure 4.25: The LIME values for PM10 VS. daily temperatures.

health risks. These results align with a meta-analysis of 29 studies employing GAM or GLM models [118].

A similar trend is observed in the city of San Diego. Figure 4.26 illustrates a comparable association between PM_{10} and temperature in assessing mortality using the LSTM model.

The interaction between air pollutants and temperature in relation to daily mortality has been extensively studied using datasets from various regions worldwide [119, 120, 121, 122]. All of these studies stratified temperature and applied Generalized Additive Models (GAMs) separately. In contrast, the LSTM-LIME method introduced in our research provides new insights into this field by enabling the in-

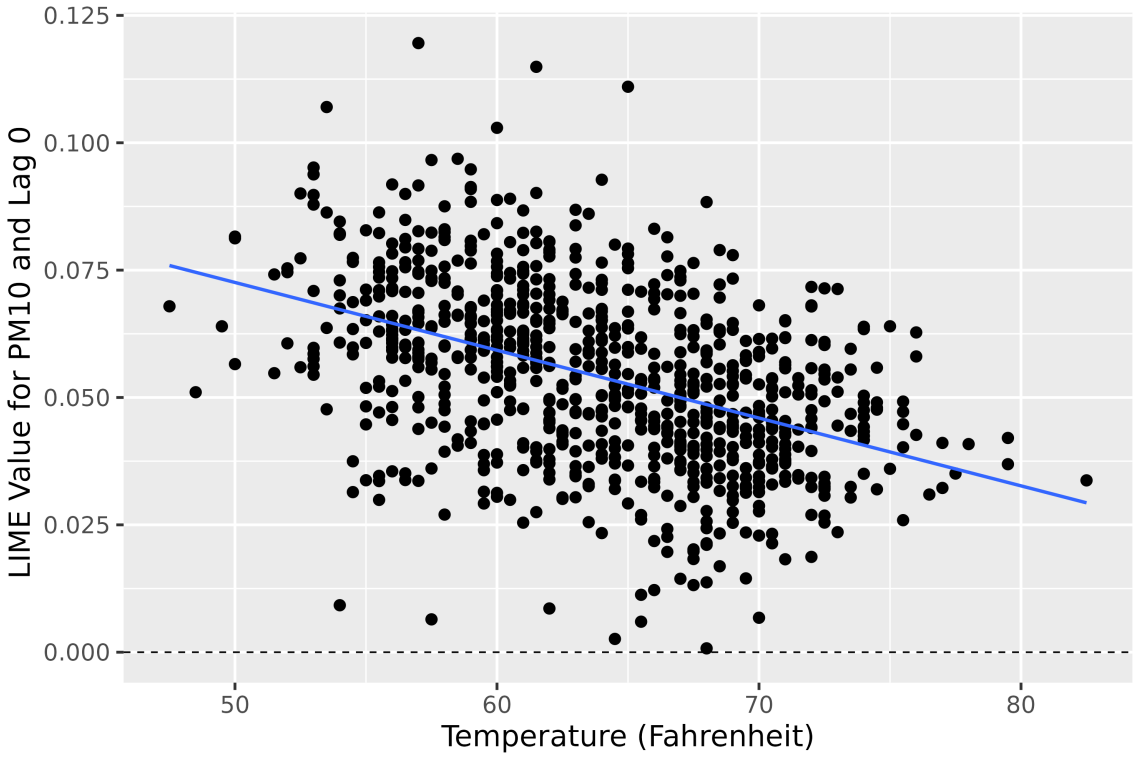

Figure 4.26: The LIME values for PM10 VS. daily temperatures with San Diego dataset.

corporation of continuous temperature variables in the analysis.

Chapter 5

Conclusions and Future Work

In this chapter, conclusions for the research in this thesis are briefly made and then future work of further improving the proposed model to facilitate its application in air pollution epidemiology is discussed.

5.1 Conclusions

In this thesis, an LSTM neural network model is developed to assess the public health outcomes of exposure to air pollution, which aims to accommodate the impacts of both distributed lags of exposure and multiple air pollutants. MIC is first used to evaluate the association between different air pollutants and between a specific air pollutant and health outcome of interest by means of standardizing information entropy-based MI. It is shown through experiments on the Toronto and Chicago datasets that the MIC can capture both the linear and nonlinear relation between different elements, and can be further used to select the most related air pollutant

(s) for a specific health outcome.

In Chapter 3, an LSTM neural network model for air pollution impacts assessment is formulated. In this formulation, an LSTM network is constructed to extract health outcome-related feature information from distributed exposure to a single or multiple air pollutant(s), and neural network layers with adaptive weights for the extracted features are developed to assess the health outcome. Experiments on the Chicago dataset from `NMMAPS` demonstrate that:

1. the LSTM model is adaptable to single and multiple air pollutant(s);

2. the LSTM model performs well in accommodating the impacts of exposure with distributed lags for both fitting and predicting the health outcome, as measured by the loss function;

3. performance of the LSTM model roughly improves with the increase of the length of exposure series, where more outcome-related temporal information is utilized (that is, higher lag);

4. compared to single pollutant, performance of the model with multiple air pollutants is better, again, as measured by the loss; and

5. the proposed model outperforms GAM in health outcome fitting accuracy, with less errors and larger R^2.

However, as mentioned in that Section, the primary issue is that we mostly use GAMs for *inference*, not just prediction! And black-box neural network models are

quite inappropriate choices for trying to understand constrained associations. This led us to Chapter 4.

The proposed LSTM model, though effective at prediction as measured by the minimization of loss, lacks interpretability for credibility verification and does not generate comparable reference coefficients. Thus, in Chapter 4 we explored the application of the LIME method to the LSTM model, using surrogate local models to interpret feature importance and evaluate their impact for the prediction of mortality. The informative LIME values for all instances across time enabled lag analysis and interaction effect analysis between air pollutants and temperature in the epidemiology models in a new sense. Analysis of portions of the NMMAPS and `AHI` datasets revealed that:

1. In the Chicago LSTM model, where PM_{10}, O_3, CO, NO_2, and temperature serve as input features, lag analysis reveals that Lag 0 and Lag 3 are the most dominant in LIME importance among all lags (0 to 4). Aggregated LIME values across these dominant lags indicate that CO has the highest influence, whereas O_3 has the lowest. In the single air pollutant model (PM_{10}), the ranking of lag importance aligns with findings from another study that utilized the traditional GAM model.

2. For the Toronto LSTM model, which incorporates NO_2, O_3, and temperature as input features, Lag 0 and Lag 4 exhibit the highest absolute LIME values among all lags (0 to 4). Across other Canadian CDs, Lag 0 consistently emerges as the most dominant in absolute LIME values, regardless of the total number

of lags considered in the input.

3. The interaction effect between temperature and air pollutants in mortality assessment models is also evident in LSTM models. Notably, cold weather is associated with increased LIME values for PM_{10} in both Houston and San Diego.

5.2 Future Work

Although the association between different elements can be evaluated comprehensively with the MIC and the proposed LSTM model shows good performance in air pollution-related health outcome prediction, there also exist limitations in their applications. These problems and several promising research opportunities that deserve attention are further discussed as follows.

1) **Redundant information.** For pairwise elements, the MIC can effectively capture their dependencies. However, for multiple air pollutants selected for some health outcome using the MIC, they may actually not be the most related factors as a whole due to the redundant health outcome-related information among them. This redundancy may weaken the strength of the selected air pollutants for outcome assessing and further raises the question of how to select the most related multi-pollutant. A promising research direction would be to select the most related candidate air pollutants for the health outcome first, and then select the least associated candidates therein.

2) **Model refinement.** Main structure parameters including number of hidden

layers and output features of the LSTM network as well as the batch size and type of the optimizer keep unchanged in previous experimentation. These hyperparameters can be further tuned for each experiment to improve the model performance. Possible parameter tuning methods include but not limited to grid search [123], randomized search [124] and Bayesian optimization [125].

3) **Distributed lags.** Historical air pollution exposure has important lagged impacts on the health outcome. Longer-term exposure sequence can provide more outcome-related features. However, if the input sequence length keeps increasing, more data noise is also brought into the LSTM model, with which the model performance may be weakened. In addition, incorporating more distributed lags increases the computational complexity and challenges the model's capability in capturing long-term dependencies. Thus, optimizing the length of input air pollution exposure sequence is a promising research direction to facilitate the application of the model to air pollution epidemiology and the methods mentioned for hyperparameter tuning may be referred to.

4) **Pooled and (or) meta analysis.** Based on the application of the LSTM model in individual studies, data or results from multiple areas can be further combined for collaborative and synthetic analyses. Either retrospective or prospective analyses can be made to provide overall assessments and more reliable results.

The innovative LSTM neural network model enhances the assessment of ambient pollutants' impact on daily mortality, demonstrating superior predictive accuracy compared to GAM models. To validate the reliability of this approach, the LIME

method is applied to LSTM models, bridging the gap between the LSTM technique and traditional GAM models in explaining the association between air pollution and mortality. LIME values are used for lag analysis, feature importance evaluation, and interaction effect analysis between temperature and air pollutants. The results from the LSTM models align with those of GAM models when tested on the same dataset, reinforcing the credibility of our proposed method. Furthermore, the LSTM-LIME approach enables quantitative analysis for integrated models with multi-pollutant and multi-lag inputs – an analysis that GAM models cannot achieve.

Bibliography

[1] Yun-Chul Hong, Jong-Tae Lee, Ho Kim, and Ho-Jang Kwon. Air pollution: a new risk factor in ischemic stroke mortality. *Stroke*, 33(9):2165–2169, 2002.

[2] Ken Donaldson, Nicholas Mills, William MacNee, Simon Robinson, and David Newby. Role of inflammation in cardiopulmonary health effects of PM. *Toxicology and Applied Pharmacology*, 207(2):483–488, 2005.

[3] Gerard Hoek, Ranjini M Krishnan, Rob Beelen, Annette Peters, Bart Ostro, Bert Brunekreef, and Joel D Kaufman. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environmental Health*, 12:1–16, 2013.

[4] Jamie I Verhoeven, Youssra Allach, Ilonca CH Vaartjes, Catharina JM Klijn, and Frank-Erik de Leeuw. Ambient air pollution and the risk of ischaemic and haemorrhagic stroke. *The Lancet Planetary Health*, 5(8):542–552, 2021.

[5] Joel Schwartz. Particulate air pollution and chronic respiratory disease. *Environmental Research*, 62(1):7–13, 1993.

[6] C Arden Pope Iii, Richard T Burnett, Michael J Thun, Eugenia E Calle, Daniel Krewski, Kazuhiko Ito, and George D Thurston. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *The Journal of the American Medical Association*, 287(9):1132–1141, 2002.

[7] Frank J Kelly. Influence of air pollution on respiratory disease. *European Medical Journal Respiratory*, 2:96–103, 2014.

[8] Douglas W Dockery, C Arden Pope, Xiping Xu, John D Spengler, James H Ware, Martha E Fay, Benjamin G Ferris Jr, and Frank E Speizer. An association between air pollution and mortality in six US cities. *New England Journal of Medicine*, 329(24):1753–1759, 1993.

[9] Jonathan M Samet, Francesca Dominici, Frank C Curriero, Ivan Coursac, and Scott L Zeger. Fine particulate air pollution and mortality in 20 US cities, 1987–1994. *New England Journal of Medicine*, 343(24):1742–1749, 2000.

[10] Paul H Fischer, Marten Marra, Caroline B Ameling, Gerard Hoek, Rob Beelen, Kees de Hoogh, Oscar Breugelmans, Hanneke Kruize, Nicole AH Janssen, and Danny Houthuijs. Air pollution and mortality in seven million adults: the dutch environmental longitudinal study (duels). *Environmental Health Perspectives*, 123(7):697–704, 2015.

[11] Jonathan M Samet, Francesca Dominici, Scott L Zeger, Joel Schwartz, and Douglas W Dockery. The national morbidity, mortality, and air pollution study. Part I: methods and methodologic issues. *Research Report (Health Effects Institute)*, 94(pt 1):5–14, 2000.

[12] Jonathan M Samet, Scott L Zeger, Francesca Dominici, Frank Curriero, Ivan Coursac, Douglas W Dockery, Joel Schwartz, and Antonella Zanobetti. The national morbidity, mortality, and air pollution study. Part II: Morbidity and mortality from air pollution in the United States. *Research Report (Health Effects Institute)*, 94(pt 2):5–79, 2000.

[13] Klea Katsouyanni, Joel Schwartz, Claudia Spix, Giota Touloumi, Denis Zmirou, Antonella Zanobetti, Bogdan Wojtyniak, JM Vonk, A Tobias, A Pönkä, S Medina, L Bachárová, and H R Anderson. Short term effects of air pollution on health: a European approach using epidemiologic time series data: the APHEA protocol. *Journal of Epidemiology and Community Health*, 50(1):12–18, 1996.

[14] Alexandros Gryparis, Bertil Forsberg, Klea Katsouyanni, Antonis Analitis, Giota Touloumi, Joel Schwartz, Evangelia Samoli, Sylvia Medina, HR Anderson, Emilia Maria Niciu, HE Wichmann, Bohumir Kriz, Mitja Kosnik, Jiri Skorkovsky, JM Vonk, and Zeynep Dörtbudak. Acute effects of ozone on mortality from the "Air Pollution and Health: A European Approach" project. *American Journal of Respiratory and Critical Care Medicine*, 170(10):1080–1087, 2004.

[15] Klea Katsouyanni and APHEA Group. APHEA project: Air pollution and health: A European approach. *Epidemiology*, 17(6):S19, 2006.

[16] Danford G Kelley. The national air pollution surveillance network in Canada. *Journal of the Air Pollution Control Association*, 29(8):794–795, 1979.

[17] Kenneth L Demerjian. A review of national monitoring networks in North America. *Atmospheric Environment*, 34(12-14):1861–1884, 2000.

[18] Francesca Dominici, Jonathan M Samet, and Scott L Zeger. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 163(3):263–302, 2000.

[19] Francesca Dominici, Aidan McDermott, Scott L Zeger, and Jonathan M Samet. On the use of generalized additive models in time-series studies of air pollution and health. *American Journal of Epidemiology*, 156(3):193–203, 2002.

[20] Richard B Davies. From cross-sectional to longitudinal analysis. *Analyzing Social and Political Change: A Casebook of Methods*, pages 20–40, 1994.

[21] Thomas Götschi, Joachim Heinrich, Jordi Sunyer, and Nino Künzli. Long-term effects of ambient air pollution on lung function: a review. *Epidemiology*, 19(5):690–701, 2008.

[22] Antonio Ciocco and Donovan J Thompson. A follow-up of Donora ten years after: methodology and findings. *American Journal of Public Health and the Nations Health*, 51(2):155–164, 1961.

[23] Lester B Lave and Eugene P Seskin. Air pollution and human health: The quantitative effect, with an estimate of the dollar benefit of pollution abatement, is considered. *Science*, 169(3947):723–733, 1970.

[24] Frederick W Lipfert. Statistical studies of mortality and air pollution: Multiple regression analyses stratified by age group. *Science of the Total Environment*, 15(2):103–122, 1980.

[25] Diane I Gibbons and Gary C McDonald. Illustrating regression diagnostics with an air pollution and mortality model. *Computational Statistics and Data Analysis*, 1:201–220, 1983.

[26] Bart D Ostro. The effects of air pollution on work loss and morbidity. *Journal of Environmental Economics and Management*, 10(4):371–382, 1983.

[27] RE Waller, AG Brooks, and MW Adler. The 1952 fog cohort study. *British Journal of Preventive and Social Medicine*, 27(1):68, 1973.

[28] Anthony T Kerigan, Charles H Goldsmith, and L David Pengelly. A three-year cohort study of the role of environmental factors in the respiratory health of children in Hamilton, Ontario: epidemiologic survey design, methods, and description of cohort. *American Review of Respiratory Disease*, 133(6):987–993, 1986.

[29] Frank E Speizer. Assessment of the epidemiological data relating lung cancer to air pollution. *Environmental Health Perspectives*, 47:33–42, 1983.

[30] Göran Pershagen and Lorenzo Simonato. Epidemiological evidence on air pollution and cancer. In *Air Pollution and Human Cancer*, pages 63–74. Springer, 1990.

[31] Kari Hemminki and Göran Pershagen. Cancer risk of air pollution: epidemiological evidence. *Environmental Health Perspectives*, 102(suppl 4):187–192, 1994.

[32] Beate Ritz, Fei Yu, Guadalupe Chapa, and Scott Fruin. Effect of air pollution on preterm birth among children born in Southern California between 1989 and 1993. *Epidemiology*, 11(5):502–511, 2000.

[33] Sung Kyun Park, Marie S O'Neill, Pantel S Vokonas, David Sparrow, and Joel Schwartz. Effects of air pollution on heart rate variability: the VA normative aging study. *Environmental Health Perspectives*, 113(3):304–309, 2005.

[34] Lie Hong Chen, Synnove F Knutsen, David Shavlik, W Lawrence Beeson, Floyd Petersen, Mark Ghamsary, and David Abbey. The association between fatal coronary heart disease and ambient particulate air pollution: are females at greater risk? *Environmental Health Perspectives*, 113(12):1723–1729, 2005.

[35] Justin Barclay, Graham Hillis, and Jon Ayres. Air pollution and the heart: cardiovascular effects and mechanisms. *Toxicological Reviews*, 24:115–123, 2005.

[36] Pierre R Band, Nhu D Le, Raymond Fang, Michele Deschamps, Andrew J Coldman, Richard P Gallagher, and Joanne Moody. Cohort study of Air Canada pilots: mortality, cancer incidence, and leukemia risk. *American Journal of Epidemiology*, 143(2):137–143, 1996.

[37] David E Abbey, Naomi Nishino, William F McDonnell, Raoul J Burchette, Synnøve F Knutsen, W Lawrence Beeson, and Jie X Yang. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. *American Journal of Respiratory and Critical Care Medicine*, 159(2):373–382, 1999.

[38] Michael Jerrett, Richard Burnett, Renjun Ma, Bruce Newbold, George Thurston, and Daniel Krewski. A cohort study of air pollution and mortality in Los Angeles. *Epidemiology*, 15(4):S46, 2004.

[39] Edward L Korn and Alice S Whittemore. Methods for analyzing panel studies of acute health effects of air pollution. *Biometrics*, pages 795–802, 1979.

[40] Alice S Whittemore and Edward L Korn. Asthma and air pollution in the Los Angeles area. *American Journal of Public Health*, 70(7):687–696, 1980.

[41] Justin L Barclay, Brian G Miller, Smita Dick, Martine Dennekamp, Isobel Ford, Graham S Hillis, Jon G Ayres, and Anthony Seaton. A panel study of air pollution in subjects with heart failure: negative results in treated patients. *Occupational and Environmental Medicine*, 66(5):325–334, 2009.

[42] Ubiratan de Paula Santos, Alfésio Luís Ferreira Braga, Dante Marcelo Artigas Giorgi, Luiz Alberto Amador Pereira, César Jose Grupi, Chin An Lin, Marcos Antonio Bussacos, Dirce Maria Trevisan Zanetta, Paulo Hilário do Nascimento Saldiva, and Mario Terra Filho. Effects of air pollution on blood pressure and heart rate variability: a panel study of vehicular traffic controllers in the city of Sao Paulo, Brazil. *European Heart Journal*, 26(2):193–200, 2005.

[43] Katharina Hildebrandt, Regina Rückerl, Wolfgang Koenig, Alexandra Schneider, Mike Pitz, Joachim Heinrich, Victor Marder, Mark Frampton, Günter Oberdörster, H Erich Wichmann, et al. Short-term effects of air pollution: a panel study of blood markers in patients with chronic pulmonary disease. *Particle and Fibre Toxicology*, 6:1–13, 2009.

[44] Nelly D Saenen, Eline B Provost, Mineke K Viaene, Charlotte Vanpoucke, Wouter Lefebvre, Karen Vrijens, Harry A Roels, and Tim S Nawrot. Recent versus chronic exposure to particulate matter air pollution in association with neurobehavioral performance in a panel study of primary schoolchildren. *Environment International*, 95:112–119, 2016.

[45] Susanna Lagorio, Francesco Forastiere, Riccardo Pistelli, Ivano Iavarone, Paola Michelozzi, Valeria Fano, Achille Marconi, Giovanni Ziemacki, and Bart D

Ostro. Air pollution and lung function among susceptible adult subjects: a panel study. *Environmental Health*, 5:1–12, 2006.

[46] Howard Kipen, David Rich, Wei Huang, Tong Zhu, Guangfa Wang, Min Hu, Shou-en Lu, Pamela Ohman-Strickland, Ping Zhu, Yuedan Wang, et al. Measurement of inflammation and oxidative stress following drastic changes in air pollution during the Beijing Olympics: a panel study approach. *Annals of the New York Academy of Sciences*, 1203(1):160–167, 2010.

[47] JJK Jaakkola. Case-crossover design in air pollution epidemiology. *European Respiratory Journal*, 21(40 suppl):81s–85s, 2003.

[48] Malcolm Maclure. The case-crossover design: a method for studying transient effects on the risk of acute events. *American Journal of Epidemiology*, 133(2):144–153, 1991.

[49] Lucas M Neas, Joel Schwartz, and Douglas Dockery. A case-crossover analysis of air pollution and mortality in Philadelphia. *Environmental Health Perspectives*, 107(8):629–631, 1999.

[50] Jong-Tae Lee and Joel Schwartz. Reanalysis of the effects of air pollution on daily mortality in Seoul, Korea: a case-crossover design. *Environmental Health Perspectives*, 107(8):633–636, 1999.

[51] Harvey Checkoway, Drew Levy, Lianne Sheppard, Joel Kaufman, Jane Koenig, and David Siscovick. A case-crossover analysis of fine particulate matter air pollution and out-of-hospital sudden cardiac arrest. *Research Report-Health Effects Institute*, 2000.

[52] Hyoung June Im, Sang Yun Lee, Ki Jung Yun, Young Su Ju, Dae Hee Kang, and Soo Hon Cho. A case-crossover study between air pollution and hospital emergency room visits by asthma attack. *Korean Journal of Occupational and Environmental Medicine*, 12(2):249–257, 2000.

[53] Daniela D'Ippoliti, Francesco Forastiere, Carla Ancona, Nera Agabiti, Danilo Fusco, Paola Michelozzi, and Carlo A Perucci. Air pollution and myocardial infarction in Rome: a case-crossover analysis. *Epidemiology*, 14(5):528–535, 2003.

[54] Francesco Forastiere, Massimo Stafoggia, Sally Picciotto, Tom Bellander, Daniela D'Ippoliti, Timo Lanki, Stephanie von Klot, Fredrik Nyberg, Pentti Paatero, Annette Peters, Juha Pekkanen, Jordi Sunyer, and CA Perucci. A case-crossover analysis of out-of-hospital coronary deaths and air pollution in Rome, Italy. *American Journal of Respiratory and Critical Care Medicine*, 172(12):1549–1555, 2005.

[55] Roger Zemek, Mieczysław Szyszkowicz, and Brian H Rowe. Air pollution and emergency department visits for otitis media: a case-crossover study in Edmonton, Canada. *Environmental Health Perspectives*, 118(11):1631–1636, 2010.

[56] Takashi Yorifuji, Etsuji Suzuki, and Saori Kashima. Hourly differences in air pollution and risk of respiratory disease in the elderly: a time-stratified case-crossover study. *Environmental Health*, 13:1–11, 2014.

[57] Cheryl S Pirozzi, Barbara E Jones, James A VanDerslice, Yue Zhang, Robert Paine III, and Nathan C Dean. Short-term air pollution and incident pneumonia. a case–crossover study. *Annals of the American Thoracic Society*, 15(4):449–459, 2018.

[58] Jerry A Hausman, Bart D Ostro, and David A Wise. Air pollution and lost work. *NBER Working Paper*, (w1263), 1984.

[59] Bart D Ostro and Susy Rothschild. Air pollution and acute respiratory morbidity: an observational study of multiple pollutants. *Environmental Research*, 50(2):238–247, 1989.

[60] Bart David Ostro. Estimating the risks of smoking, air pollution, and passive smoke on acute respiratory conditions. *Risk Analysis*, 9(2):189–196, 1989.

[61] Toshiro Tango. Effect of air pollution on lung cancer: a Poisson regression model based on vital statistics. *Environmental Health Perspectives*, 102(suppl 8):41–45, 1994.

[62] Joel Schwartz. Particulate air pollution and daily mortality in Detroit. *Environmental Research*, 56(2):204–213, 1991.

[63] Arnoud P Verhoeff, Gerard Hoek, Joel Schwartz, and Joop H van Wijnen. Air pollution and daily mortality in Amsterdam. *Epidemiology*, 7(3):225–230, 1996.

[64] Lianne Sheppard and Doris Damian. Estimating short-term pm effects accounting for surrogate exposure measurements from ambient monitors. *Environmetrics: The official journal of the International Environmetrics Society*, 11(6):675–687, 2000.

[65] Claudia Spix, H Ross Anderson, Joel Schwartz, Maria Angela Vigotti, Alain Letertre, Judith M Vonk, Giota Touloumi, Franck Balducci, Tomasz Piekarski, Ljuba Bacharova, Aurelio Tobias, antti Pönkä, and Klea Katsouyanni. Short-term effects of air pollution on hospital admissions of respiratory diseases in Europe: a quantitative summary of APHEA study results. *Archives of Environmental Health: An International Journal*, 53(1):54–64, 1998.

[66] Klea Katsouyanni, Giota Touloumi, Evangelia Samoli, Alexandros Gryparis, Alain Le Tertre, Yannis Monopolis, Giuseppe Rossi, Denis Zmirou, Ferran Ballester, Azedine Boumghar, Hugh Ross Anderson, Bogdan Wojtyniak, Anna Paldy, Rony Braunstein, Juha Pekkanen, Christian11 Schindler, and Joel Schwartz. Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. *Epidemiology*, 12(5):521–531, 2001.

[67] Ariana Zeka and Joel Schwartz. Estimating the independent effects of multiple pollutants in the presence of measurement error: an application of a

measurement-error–resistant technique. *Environmental Health Perspectives*, 112(17):1686–1690, 2004.

[68] Antonella Zanobetti, Matt P Wand, Joel Schwartz, and Louise M Ryan. Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics*, 1(3):279–292, 2000.

[69] Francesca Dominici, Aidan McDermott, Scott L Zeger, and Jonathan M Samet. Airborne particulate matter and mortality: timescale effects in four US cities. *American Journal of Epidemiology*, 157(12):1055–1065, 2003.

[70] Wesley S Burr, Hwashin H Shin, and Glen Takahara. Synthetically lagged models. *Statistics and Probability Letters*, 144:37–43, 2019.

[71] Guowen Huang, Marta Blangiardo, Patrick E Brown, and Monica Pirani. Long-term exposure to air pollution and COVID-19 incidence: a multi-country study. *Spatial and Spatio-temporal Epidemiology*, 39:100443, 2021.

[72] Wesley S Burr, Glen Takahara, and Hwashin H Shin. Bias correction in estimation of public health risk attributable to short-term air pollution exposure. *Environmetrics*, 26(4):298–311, 2015.

[73] Dimitris Evangelopoulos, Klea Katsouyanni, Joel Schwartz, and Heather Walton. Quantifying the short-term effects of air pollution on health in the presence of exposure measurement error: a simulation study of multi-pollutant model results. *Environmental Health*, 20:1–13, 2021.

[74] Francesca Dominici, Antonella Zanobetti, Joel Schwartz, Danielle Braun, Ben Sabath, and Xiao Wu. Assessing adverse health effects of long-term exposure to low levels of ambient air pollution: implementation of causal inference methods. *Research Reports: Health Effects Institute*, 2022, 2022.

[75] Hwashin Hyun Shin, Rajendra Prasad Parajuli, Aubrey Maquiling, and Marc Smith-Doiron. Temporal trends in associations between ozone and circulatory

mortality in age and sex in Canada during 1984–2012. *Science of the Total Environment*, 724:137944, 2020.

[76] Hwashin H Shin, James Owen, Anna O Delic, Muzeyyen Kabasakal, and Stéphane Buteau. Modifying factors and temporal trends of adverse health effects of short-term exposure to pm2.5 in Canada (2001–2018). *Science of The Total Environment*, 955:177046, 2024.

[77] Hwashin Hyun Shin, James Owen, Aubrey Maquiling, Rajendra Prasad Parajuli, and Marc Smith-Doiron. Circulatory health risks from additive multi-pollutant models: short-term exposure to three common air pollutants in canada. *Environmental Science and Pollution Research*, 30(6):15740–15755, 2023.

[78] Huawei Han and Wesley S. Burr. Mutual Information Based Correlation Analysis of Health-Related Multiple Air Pollutants. In *Proceedings of the 6th International Conference on Statistics: Theory and Applications*, August 2024.

[79] Government of Canada Open Data Portal. National Air Pollution Surveillance Program. https://data-donnees.az.ec.gc.ca/data/air/monitor/national-air-pollution-surveillance-naps-program/?lang=en, 2024.

[80] Government of Canada Open Data Portal. National Air Pollution Surveillance Program. https://open.canada.ca/data/en/dataset/1b36a356-defd-4813-acea-47bc3abd859b, 2024.

[81] Government of Canada Open Data Portal. Historical Climate Data. https://climate.weather.gc.ca/, 2024.

[82] Roger D Peng and Francesca Dominici. Statistical methods for environmental epidemiology with R. *R: a case study in air pollution and health*, 2008.

[83] Roger D Peng and Leah J Welty. The NMMAPS data package. *R news*, 4(2):10–14, 2004.

[84] National Weather Service. National Oceanic and Atmospheric Administration Online Weather Data. https://www.weather.gov/wrh/Climate?wfo=lot, 2024.

[85] Shannon Jarvis. Particulate matter component analyses in relation to public health in Canada. Master's thesis, Trent University (Canada), 2023.

[86] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 69(6):066138, 2004.

[87] Claude Elwood Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[88] David N Reshef, Yakir A Reshef, Hilary K Finucane, Sharon R Grossman, Gilean McVean, Peter J Turnbaugh, Eric S Lander, Michael Mitzenmacher, and Pardis C Sabeti. Detecting novel associations in large data sets. *Science*, 334(6062):1518–1524, 2011.

[89] David N Reshef, Yakir A Reshef, Michael Mitzenmacher, and Pardis C Sabeti. Cleaning up the record on the maximal information coefficient and equitability. *Proceedings of the National Academy of Sciences*, 111(33):E3362–E3363, 2014.

[90] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[91] Davide Albanese, Michele Filosi, Roberto Visintainer, Samantha Riccadonna, Giuseppe Jurman, and Cesare Furlanello. Minerva and minepy: a C engine for the MINE suite and its R, Python and MATLAB wrappers. *Bioinformatics*, 29(3):407–408, 2013.

[92] Huawei Han and Wesley S. Burr. Development of LSTM Based Model for Air Pollutants-Related Public Health Consequence Assessment. In *Proceedings of the 6th International Conference on Statistics: Theory and Applications*, 2024.

[93] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[94] MI Jordan. Serial order: a parallel distributed processing approach. Technical report, California University, San Diego, La Jolla (USA). Inst. for Cognitive Science, 1986.

[95] Jeffrey L Elman. Finding structure in time. *Cognitive Science*, 14(2):179–211, 1990.

[96] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.

[97] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166, 1994.

[98] Barak A Pearlmutter. Gradient calculations for dynamic recurrent neural networks: A survey. *IEEE Transactions on Neural networks*, 6(5):1212–1228, 1995.

[99] Kanchan M Tarwani and Swathi Edem. Survey on recurrent neural network in natural language processing. *International Journal of Engineering Trends and Technology*, 48(6):301–304, 2017.

[100] Alex Graves, Navdeep Jaitly, and Abdel-rahman Mohamed. Hybrid speech recognition with deep bidirectional LSTM. In *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, pages 273–278. IEEE, 2013.

[101] Shashi Pal Singh, Ajai Kumar, Hemant Darbari, Lenali Singh, Anshika Rastogi, and Shikha Jain. Machine translation using deep learning: An overview. In *2017 International Conference on Computer, Communications and Electronics (comptelix)*, pages 162–167. IEEE, 2017.

[102] Brian S Freeman, Graham Taylor, Bahram Gharabaghi, and Jesse Thé. Forecasting air quality time series using deep learning. *Journal of the Air and Waste Management Association*, 68(8):866–886, 2018.

[103] Pantelis Linardatos, Vasilis Papastefanopoulos, and Sotiris Kotsiantis. Explainable AI: A review of machine learning interpretability methods. *Entropy*, 23(1):18, 2020.

[104] Wojciech Samek, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, and Klaus-Robert Müller. *Explainable AI: Interpreting, explaining and visualizing deep learning*, volume 11700. Springer Nature, 2019.

[105] Nadia Burkart and Marco F Huber. A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, 70:245–317, 2021.

[106] Jianlong Zhou, Amir H Gandomi, Fang Chen, and Andreas Holzinger. Evaluating the quality of machine learning explanations: A survey on methods and metrics. *Electronics*, 10(5):593, 2021.

[107] Wojciech Samek, Grégoire Montavon, Sebastian Lapuschkin, Christopher J Anders, and Klaus-Robert Müller. Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3):247–278, 2021.

[108] Julia Amann, Alessandro Blasimme, Effy Vayena, Dietmar Frey, and Vince I Madai. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC medical informatics and decision making*, 20:1–9, 2020.

[109] Bryan Lim, Sercan Ö Arık, Nicolas Loeff, and Tomas Pfister. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4):1748–1764, 2021.

[110] Chaojie Li, Zhaoyang Dong, Lan Ding, Henry Petersen, Zihang Qiu, Guo Chen, and Deo Prasad. Interpretable memristive LSTM network design for probabilistic residential load forecasting. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 69(6):2297–2310, 2022.

[111] Leon Chua. Memristor–the missing circuit element. *IEEE Transactions on circuit theory*, 18(5):507–519, 1971.

[112] Xin Jing, Jungang Luo, Ganggang Zuo, and Xue Yang. Interpreting runoff forecasting of long short-term memory network: an investigation using the integrated gradient method on runoff data from the Han River basin. *Journal of Hydrology: Regional Studies*, 50:101549, 2023.

[113] Aum Patil, Amey Wadekar, Tanishq Gupta, Rohit Vijan, and Faruk Kazi. Explainable LSTM model for anomaly detection in HDFS log file using layerwise relevance propagation. In *2019 IEEE Bombay Section Signature Conference (IBSSC)*, pages 1–6. IEEE, 2019.

[114] Guannan Li, Fan Li, Chengliang Xu, and Xi Fang. A spatial-temporal layerwise relevance propagation method for improving interpretability and prediction accuracy of LSTM building energy prediction. *Energy and Buildings*, 271:112317, 2022.

[115] Abdulmuneem Bashaiwth, Hamad Binsalleeh, and Basil AsSadhan. An ex-

planation of the LSTM model used for DDoS attacks classification. *Applied Sciences*, 13(15):8820, 2023.

[116] Enbin Yang, Hao Zhang, Xinsheng Guo, Zinan Zang, Zhen Liu, and Yuanning Liu. A multivariate multi-step LSTM forecasting model for tuberculosis incidence with model explanation in Liaoning province, China. *BMC Infectious Diseases*, 22(1):490, 2022.

[117] Eyal Winter. The Shapley value. *Handbook of game theory with economic applications*, 3:2025–2054, 2002.

[118] Fei Chen, Zhiwei Fan, Zhijiao Qiao, Yan Cui, Meixia Zhang, Xing Zhao, and Xiaosong Li. Does temperature modify the effect of PM10 on mortality? A systematic review and meta-analysis. *Environmental Pollution*, 224:326–335, 2017.

[119] Kai Chen, Kathrin Wolf, Susanne Breitner, Antonio Gasparrini, Massimo Stafoggia, Evangelia Samoli, Zorana Jovanovic Andersen, Getahun Bero-Bedada, Tom Bellander, Frauke Hennig, Bénédicte Jacquemin, Juha Pekkanen, Regina Hampel, Josef Cyrys, Annette Peters, and Alexandra Schneider. Two-way effect modifications of air pollution and air temperature on total natural and cardiovascular mortality in eight european urban areas. *Environment International*, 116:186–196, 2018.

[120] Steven Roberts. Interactions between particulate air pollution and temperature in air pollution mortality time series studies. *Environmental Research*, 96(3):328–337, 2004.

[121] Rennie Xinrui Qin, Changchun Xiao, Yibin Zhu, Jing Li, Jun Yang, Shaohua Gu, Junrui Xia, Bin Su, Qiyong Liu, and Alistair Woodward. The interactive effects between high temperature and air pollution on mortality: A time-series analysis in hefei, china. *Science of The Total Environment*, 575:1530–1537, 2017.

[122] Yuexin Cheng and Haidong Kan. Effect of the interaction between outdoor air pollution and extreme temperature on daily mortality in shanghai, china. *Journal of Epidemiology*, 22(1):28–36, 2012.

[123] Hussain Alibrahim and Simone A Ludwig. Hyperparameter optimization: Comparing genetic algorithm against grid search and bayesian optimization. In *2021 IEEE Congress on Evolutionary Computation (CEC)*, pages 1551–1559. IEEE, 2021.

[124] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(2), 2012.

[125] A Helen Victoria and Ganesh Maragatham. Automatic tuning of hyperparameters using Bayesian optimization. *Evolving Systems*, 12(1):217–223, 2021.

APPENDICES

Functions Used in LSTM network

The sigmoid, softmax and tanh functions used in the LSTM model is as Eqns. (5.1-5.3).

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{5.1}$$

$$SF(i) = \frac{e^{i}}{\sum_{j} e^{j}} \tag{5.2}$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{5.3}$$