Machine Learning Based Calibration Techniques for Low-Cost Air Quality Sensors

Mohammad Sharafat Ali

Electronic and Computer Engineering

Massey University

Thesis

For

Doctor of Philosophy

May 2024

ACKNOWLEDGEMENT

First and foremost, I convey my deepest gratitude to the Almighty, the most merciful and gracious, for giving me the guidance, patience, knowledge, and determination to complete this research. This thesis is the result of the work where many people have accompanied and supported me. I now have the opportunity to express my gratitude to all of them. I want to convey my sincerest gratitude, regards and thanks to my supervisors, Professor Fakhrul Alam, Dr. Khalid Arif and Professor Johan Potgieter, for their support and guidance. I am especially indebted to Professor Alam whose mentorship has been instrumental throughout this doctoral study. As my supervisor, he has taught me more than I could ever give him credit for. It was my great pleasure to be a part of the Department of Mechanical and Electrical Engineering, School of Food and Advanced Technology and Massey University. I am also grateful for the financial support in the form of a doctoral scholarship offered by the New Zealand Product Accelerator (NZPA). I thank Auckland Council and National Institute of Water and Atmospheric Research (NIWA) for their support, as well as all the members of the CAIRNet team for their assistance with the sensor development. I would also like to thank all my fellow researchers, faculty members and administrators for their heartiest cooperation. I also acknowledge Dr. Nasim Ahmed and Dr. Daniel Konings for their valuable time and suggestions. I thank De Vito *et al.* and Liang *et al.* for making their research data available for this study. Finally, I thank my family for always being there as an inspiration and for their continuous support and encouragement throughout all my studies and work. My late father Dr. Maksud Ali, who was an excellent researcher in his own right, is and will always be my inspiration. I want to thank my mother, Shirin Ahmed, whose love, wisdom, and guidance are with me in whatever I pursue. Most importantly, I wish to thank my loving and supportive wife, Fatima Zohora, for providing me with unending inspiration in the pursuit of this doctoral study.

ABSTRACT

Breathable air is the single most essential element for life on earth. Polluted air, contaminated by particulate matter and harmful gases, poses numerous risks to health and the environment, especially in urban areas with large populations and many active sources of air pollution. Therefore, researchers from a wide range of disciplines have been working on mitigating the impact of air pollution. Monitoring ambient air pollution is one of the means to ensure public health safety, raise public awareness and build a sustainable urban environment. However, conventional air quality monitoring stations are mostly confined to a few locations due to their costly equipment and large sizes. As a result, although these monitoring stations provide accurate air pollution data, they can only offer a low-fidelity picture of air quality in a large city, leading to a poor spatial resolution of urban pollution data. Low-cost sensor (LCS) technologies aim to address this challenge and intend to make it possible to monitor air quality at a high spatio-temporal resolution. The pollutant data captured by these LCSs are less accurate than their conventional counterparts and thus require calibration techniques to improve their accuracy and reliability. Researchers have proposed different calibration methods and techniques to improve the accuracy of the LCSs, including machine learning based calibration models. This thesis investigates and proposes several machine learning-based techniques for calibrating low-cost ambient gas sensors and rigorously benchmarks their performance using a robust training, validation and testing method. Based on the findings, One Dimensional Convolutional Neural Network (1DCNN) and Gradient Boosting Regression (GBR) based calibration techniques provide consistently accurate performance. Both of these machine learning techniques, which have not been widely used or evaluated for low-cost ambient gas sensor calibration, can improve the state of the art. This research also demonstrates that readily available and previously unemployed co-variate data, namely the number of days the sensor has been deployed and the time of day at which the reading is taken, can significantly improve the accuracy of Machine Learning based calibration algorithms.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	*Definition*
1DCNN	One Dimensional Convolutional Neural Network
AB	Adaptive Boosting
AE	Auxiliary Electrode
AQI	Air Quality Index
AQM	Air Quality Monitoring
ANN	Artificial neural networks
BLR	Bayesian Linear Regression
CNN	Convolution Neural Network
CRMSE	Centered Root Mean Squared Error
EC	Electrochemical
EGB	Extreme Gradient Boosting
IoT	Internet of Things
GBR	Gradient Boosting Regression
GPR	Gaussian Processes Regressors
GRU	Gated Recurrent Unit
HDMR	High-Dimensional Model Representation
KNN	K-Nearest Neighbors
LCS	Low-Cost Sensors
LOD	Limit of Detection
LSTM	Long Short-Term Memory
LR	Linear Regression
MAE	Mean Absolute Error
MBE	Mean Bias Error
ML	Machine Learning
MLP	Multi-layer Perceptron

MLR	Multiple Linear Regression
MOS	Metal-oxide-semiconductor
MOX	Metal Oxide
MSE	Mean Squared Error
nCRMSE	Normalized Centered Root Mean Squared Error
nMBE	Normalized Mean Bias Error
nRMSE	Normalized Root Mean Squared Error
NMHC	Non-methanic Hydrocarbons
NN	Neural Networks
PM	Particulate Matter
PR	Polynomial Regression
PPB	Parts Per Billion
PPM	parts per Million
QR	Quadratic Regression
RC	Reservoir Computing
RH	Relative Humidity
RFR	Random Forest Regression
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Networks
SD	Standard Deviation
SI	International System of Units
SNR	Signal to Noise Ratio
STP	Standard Temperature and Pressure
SVR	Support Vector Regression
T	Temperature
TTS1	Test Train Split 1
TTS2	Test Train Split 2

VOC	Volatile Organic Compound
WE	Working Electrode
WHO	World Health Organization

CHAPTER 1

INTRODUCTION

1.1 Introduction

Air pollution in urban areas adversely affects public health, environment, and quality of life. A large part of the world's population currently lives in regions where air pollution levels have exceeded the specified limits of the World Health Organization (WHO) [1, 2]. Ambient air pollution has been associated with many serious health issues, including cancer, respiratory infections and cardiopulmonary diseases [3-6]. Monitoring ambient air pollution can be an effective tool for governments, policymakers and scientists working to alleviate this issue.

Conventional air quality monitoring systems are expensive, bulky and stationary, confining monitoring to only a select number of spots [7]. However, the pollutant concentration in ambient air can vary significantly within a few hundred meters. Therefore, the sparse coverage provided by conventional monitoring systems results in a lower spatial resolution, which fails to provide sufficient real-time data to quantify air pollution in a large area [8]. While the data from these stations are accurate, the poor spatial resolution hinders the generation of robust, city-wide air quality data. Therefore, monitoring air pollution levels over a large area using traditional monitoring systems is very difficult to accomplish [9]. Low-Cost Sensors (LCS) have been identified as an option to supplement the information captured by conventional air quality monitoring systems [10-14]. Researchers have proposed to interpolate the information gaps between these sparsely distributed air quality monitoring stations with LCS to increase the spatio-temporal resolution [9, 15]. Thus, researchers have started proposing an extensive network of sensor nodes consisting of inexpensive instrumentation to increase the spatio-temporal resolution of the present air quality monitoring [10, 16, 17]. Many countries worldwide have started to adopt this approach to monitor urban pollutants at high spatial resolution [10, 14, 18].

Real-time data of air pollution at a large number of locations may also raise the general public's awareness regarding their health and could lead to the uptake of

sustainable living and green technology, better driving approaches and reduced air pollution [10, 19]. Researchers have developed LCS and other relevant supporting technologies in the past few decades that can be used to gather real-time air pollution data [18, 20-27]. Such sensors have been used for pollutant measurements in various scenarios such as road-side pollution measurement [25, 28, 29], rural and urban air pollution measurement [30-37], mobile vehicular pollution measurement [38-40], source attribution [41], monitoring personal exposure [11, 42], etc. Several review reports have addressed many aspects of LCS technologies in the last decade and summarized the development of the LCS hardware, network, and calibration techniques [9, 11, 21, 43-47].

One of the main drawbacks of the LCS is that they are usually influenced by ambient factors more than the conventional sensors [9, 35, 45, 46] and as a result, they are comparatively less accurate. The accuracy of these sensors seems to vary significantly due to many factors including the calibration methods used [18, 48-50]. Although, previous studies indicate that LCS are not a replacement for conventional monitoring equipment and may be suitable for screening and measuring pollution patterns through small areas [9], sophisticated calibration may allow us to accomplish more than this. Calibration can facilitate the deployment of the LCS by ensuring that the gap between the results from LCS and expensive sensors is narrowed down [34, 51, 52]. Research is ongoing regarding the real-world implementation of low-cost air quality monitoring sensors [14, 16, 18, 31, 53]. However, many aspects and features are yet to be determined for calibrating LCS to higher accuracy.

Many innovative methods have been proposed to improve the accuracy and the operability of the LCS and researchers have been working to develop calibration strategies and techniques [21, 43, 44] . These sensors can be calibrated by co-locating them with accurate sensors so that the calibrated measurements of the LCS closely agree with the co-located accurate reference sensor [21]. The co-located measurements are often performed during “field deployment” as it is difficult to emulate the inherently complex nature of the ambient conditions in a controlled lab setup [43, 54]. The key aspect of the calibration is the training of regression models to capture the complex, often non-linear, relationship between the raw sensor output and the ground truth provided by the accurate, reference sensor. Different calibration methods have been proposed by researchers, with recent literature showing a growing popularity of machine learning

based techniques. However, there is no 'one-for-all' machine learning technique that can calibrated the LCS up to the highest level of performance for every real-world scenario.

While research on low-cost air pollution monitoring sensors dates back to many years, recent advancements in the field of wireless communication, calibration techniques and internet based applications have managed to attract the interest of many researchers [21-24]. Spinelle *et al.* [30, 31, 46, 55, 56] and Esposito *et al.* [50, 57, 58] studied different setups and methods to calibrate LCS nodes with multiple pollutant and meteorological sensors and found a strong correlation between the LCS response and meteorological parameters such as temperature and relative humidity as well as cross-sensitive gases. De Vito *et al.* [10, 16, 17] reported similar correlations and used different calibration techniques to overcome the errors caused by them. Many research groups have proposed different computational approaches of calibration to improve the accuracy of LCS in the past. Among these, classic statistical regressions such as Linear Regression (LR) and Multiple Linear Regression (MLR) are still being employed in recent works [10, 14, 18, 19, 53, 59-64]. State of the art calibration methods include supervised Machine Learning (ML) techniques such as Support Vector Regression (SVR) [17, 57, 58, 65-67], ensemble ML techniques like Random Forest Regression (RFR) [34, 65, 67-73] and Neural Networks (NN) such as Multilayer Perceptron (MLP) [10, 16, 17, 50, 53, 57, 58, 71, 74]) and Recurrent Neural Networks (RNN) [17, 57, 58, 68, 75-77] also have been proposed by various researchers. However, many of the low-cost air pollution monitoring systems reported in the literature do not follow a rigorous method for selecting suitable calibration processes for their developed LCS prototypes. With a few exceptions [17, 18], most of these studies typically utilize only one set of data to demonstrate the calibration performance, making it difficult to ascertain the generalizability of the techniques. With the rapid adoption of Internet of Things (IoT) technology [78] taking place in many sectors, it is high time to understand the opportunities of low-cost air quality monitoring systems in modern technology and challenges to overcome towards the next generation air quality monitoring system.

1.2 Problem Statement

Literature review on air quality monitors, LCS and next generation air quality monitoring systems bring forth the following questions to attention:

1. Can advanced machine learning techniques improve the performance of LCS?
2. Which algorithms provide good accuracy on a consistent basis for LCS calibration?
3. What inputs or co-variate factors need to be considered to improve the performance of the algorithms during low-cost air pollutant sensor calibration?

By answering these research questions, this research aims to suggest suitable calibration models and guidelines for low-cost air quality sensors.

1.3 Scope of Work

This work utilizes existing datasets from research groups working with field deployment of low-cost air quality monitoring. The LCS used in these field deployments were assembled by the research groups using commercially available components. The said research groups also collected the pollutant data. Only datasets reported in high-impact journal articles were selected in this PhD work to ensure the quality of the data. These datasets included ground truth data from co-located reference sensor to develop and evaluate different calibration techniques. Existing machine learning techniques have been adopted and modified through a rigorous training-validation-testing process using these collected datasets to make recommendations for low-cost air quality sensor calibration.

1.4 Organization of the Thesis

This thesis has been compiled with six chapters. The first chapter is introduction where the overall work is briefly introduced followed by problem statements and scope of this work. Chapter 2 presents the literature review relevant to this research covering the details of air pollution, low-cost sensors, sensor calibration and summary of prior works reported by researchers. The next chapter (Chapter 3) describes the methodology employed for this research. The various aspects of the gas pollutant datasets that were

utilized for the study are described in this chapter. This is followed by a detailed, step by step description of the calibration process. The proposed calibration algorithms and performance metrics that are used to evaluate the efficacy of the various algorithms are also described in this chapter. The results are presented in Chapters 4 and 5. Chapter 4 focuses on benchmarking various ML-based calibration techniques and identification of the algorithms that can provide consistently good accuracy for LCS calibrations. Chapter 5 introduces co-variate factors that can be exploited to improve the performance of the algorithms. Chapter 6 concludes this thesis and provides some suggestions for future work.

CHAPTER 2

LITERATURE REVIEW

This chapter introduces air pollutants, different terminologies relevant to air pollution, commonly used sensors for measuring ambient pollutants, modern air quality monitoring systems, and different calibration methods.

2.1 Air Pollutants

The ambient air is a combination of mainly oxygen and nitrogen. However, smaller amounts of other components can be found in the air surrounding us, some of which are considered harmful pollutants. Table 2.1 shows the composition of a typical sample of dry air. However it should be mentioned that the presented composition slightly varies in ambient air with location and wind speed [7]. However, even if the approximate atmospheric concentrations show no noticeable changes, a moderate increase in pollutant concentration levels can compromise human health. For example, a CO concentration of 800 ppm in the air causes a headache, dizziness and nausea, while an increased concentration of 12800 ppm causes death within three minutes of exposure [79]. Another significant air pollutant is NO_2 and breathing in NO_2 can be associated with a range of health impacts including increased susceptibility to asthma and respiratory illness. Inhaling a higher level of NO_2 (over 100 ppb) for a long period of time can also increase the risk of premature death [80].

Another significant air pollutant is the Particulate Matter (PM); the PM concentration is considered a common indicator of air pollution [81]. PM is the combination of a complex mixture of solid and liquid particles of organic and inorganic substances including sulphate, sodium chloride, ammonia, black carbon, nitrates, mineral dust, and water [81, 82]. Chronic exposure to PM can lead to respiratory and cardiovascular diseases, and lung cancer [83, 84]. PMs are typically classified into three general categories by their size in diameter. PM_{10} describes the solid and fluid air particles with a diameter smaller than 10 µm, $PM_{2.5}$ diameter is smaller than 2.5 µm and ultrafine particles ($PM_{0.1}$) are nano-particles with diameters below 0.1 µm [21, 82, 85, 86].

However, the PM measurement and low-cost PM sensor calibration take a slightly different route from the low-cost gas sensors. As this research focuses on the latter, PM sensor calibration is not explored in this thesis.

Table 2.1: Composition of dry air.

Substance	**Proportion by volume (%)**
Nitrogen (N_2)	78.1
Oxygen (O_2)	20.9
Argon (Ar)	0.93
Carbon dioxide (CO_2)	0.035
Neon (Ne)	0.0018
Ozone (O_3)	0.0008
Helium (He)	0.00052
Methane (CH_4)	0.00017
Hydrogen (H_2)	0.000053
Nitrous oxide (N_2O)	0.000031
Carbon monoxide (CO)	0.000025
Sulfur dioxide (SO_2)	0.00001
Xenon (Xe)	0.0000087
Nitrogen dioxide (NO_2)	0.000002
Ammonia (NH_3)	0.0000003

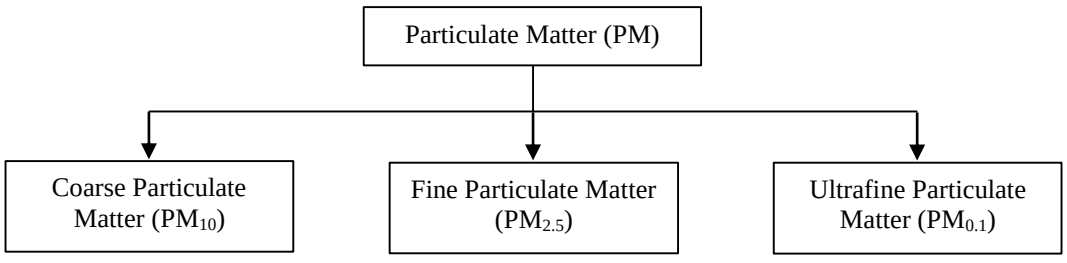


Figure 2.1: Classification of PM according to their sizes.

The five major pollutants typically reported to specify air quality are PM, ground-level O_3, CO, SO_2 and NO_2 [81, 85, 87-90]. WHO has released thresholds for health-

harmful pollution levels (as an annual or hours-based mean concentration) for these pollutants as guideline values [81]. Table 2.2 describes the principal sources and health effects of these pollutants along with their guideline values provided by WHO [1, 80, 91]. Although, all these air pollutants have adverse effect to the human health, some are more dangerous than others depending on factors like chances of getting exposed to these pollutants, duration of exposer, quantity of pollutant during exposer [80] etc.

Table 2.2: Sources, health effects and guideline values for major air pollutants.

Pollutant	**Principle Sources**	**Health Effects**	**Guideline Values**
Particulate Matter (PM)	Microscopic metals, compounds, mineral dust, water etc.	PM_{10} can penetrate and stay inside the lungs and $PM_{2.5}$ can further penetrate the lung barrier and enter the human blood system. Both causes cardiovascular diseases, respiratory failure and lung cancer.	50 $\mu g/m^3$ (24 hours) for PM_{10} and 25 $\mu g/m^3$ (24 hours) for $PM_{2.5}$ 20 $\mu g/m^3$ (1 year) for PM_{10} and 10 $\mu g/m^3$ (1 year) for $PM_{2.5}$
Ozone (O_3)	Formed by the photochemical reaction of nitrogen oxides and volatile organic compounds (VOCs) from vehicle and industry emissions	Causes breathing problems like asthma and reduces lung functionality leading towards many lung diseases	100 $\mu g/m^3$ (~ 50 ppb) for 8 hours
Carbon monoxide (CO)	Fuel combustion, heating, tobacco smoke, chimneys and furnaces etc.	Triggers fatigue and chest pain in low concentrations, and causes headaches, dizziness, impaired vision, nausea in high concentrations. Also can be fatal at very high level of concentrations.	10 mg/m^3 (~ 8.7 ppm) for 8 hours, 30 mg/m^3 (~ 26.2 ppm) for 1 hour, 60 mg/m^3 (~ 52.4 ppm) for 30 minutes and 100 mg/m^3 (~ 87.3 ppm) for 15 minutes
Sulphur dioxide (SO_2)	Produced from the burning of fossil fuels and sulphur-containing minerals	Causes respiratory infections, coughing, mucus secretion, asthma and chronic bronchitis	500 $\mu g/m^3$ (~ 190 ppb) for 10 minutes and 20 (~ 8 ppb) $\mu g/m^3$ for 24 hours

Nitrogen dioxide (NO_2)	Emitted from combustion, power generation, and engines in vehicles and ships	Reduces lung functionality and increases chances of bronchitis in asthmatic children	200 µg/m^3 (~ 106 ppb) for 1 hour and 40 µg/m^3 (~ 21 ppb) for 1 year

2.2 Air Pollutant Concentration Measurement Units

Generally, two different concentration unit sets are used, namely, volumetric units and gravimetric units [7]. Volumetric units state the ratio of the number of pollutant gas molecules to the total number of air molecules which is equivalent to the mixing ratio between the volume of the pollutant and the original air. Gravimetric units specify the mass of the pollutant material per unit volume of air. Volumetric units are invariant with changing temperature and pressure. Therefore, volumetric units like ppm (parts per million, 10^{-6}), ppb (parts per billion, 10^{-9}) and ppt (parts per trillion, 10^{-12}) have been used by most of the researchers. However, both gravimetric and volumetric systems are in use while measuring pollutant concentrations and, it is reasonably simple to convert one-unit system to another. This conversion factor is based on the gas's molecular weight (M). For example, the gas concentration in $X(mg/m^3)$ and $Y(ppm)$ at STP (Standard Temperature and Pressure) will be:

$$X\ (mg/m^3) = 0.0409 \times Y(ppm) \times M \quad 2.1$$

$$Y(ppm) = 24.45 \times X(mg/m^3) \div M \quad 2.2$$

The relationship for the typical concentrations for ambient pollutants between these two unit sets is summarized in Table 2.3.

Table 2.3: Relationship between volumetric and gravimetric units.

Volumetric		**Gravimetric**	
Expression	***Order***	***Expression***	***SI Expression***
ppm (parts per million)	10^{-6}	mg m^{-3}	µmol mol^{-1}
ppb (parts per billion)	10^{-9}	µg m^{-3}	nmol mol^{-1}
ppt (parts per trillion)	10^{-12}	ng m^{-3}	pmol mol^{-1}

2.3 Low-Cost Sensors (LCS)

At present, ambient air quality monitoring systems are mainly confined to only a small number of locations due to their high cost, resulting in a limited observation of real-time air quality index over a larger area [92]. An intuitive approach to tackle this issue could be to increase the spatiotemporal resolution of available air pollution monitoring equipment. A driving factor that enables the increased monitoring efforts is the development of air pollution sensors with lower cost [9, 46, 47, 93].

2.3.1 Properties of Air Quality Sensors

Suitability of the sensors for measuring the concentration of pollutants in the air depends on a few important properties such as specificity [94], sensitivity [95], reliability [96], reproducibility [95], stability [97], response time [17], accuracy and precision [98], and cost [35, 46]. Table 2.4 describes these properties along with the corresponding features of both accurate but expensive sensors and LCS.

LCS tend to be affected by environmental factors and may suffer from relatively low sensitivity and specificity [20, 45, 94, 95]. For obvious reasons, LCS use cheaper hardware than their more accurate counterparts. As a result, environmental factors such as temperature and relative humidity can easily affect the LCS response towards a pollutant [30, 31]. Also, these cheap components efficiently respond to a range of gases instead of being specific towards a target pollutant gas [99]. LCS can also drift over time and have poor reliability and stability [100-103]. The low-cost components have a short usable lifespan and thus LCS are well known for responding much differently to an air pollutant after some time in their deployment [100-103].

Some research groups are focusing on developing affordable but accurate LCS and pollutant measurement systems while other research groups are trying to develop new calibration techniques to improve the accuracy of currently available LCS [44, 45, 47, 104]. Our focus of this research work is on the calibration techniques for available LCS. A condensed summary of LCS hardware and pollutant measuring technologies is compiled in Section 2.3.2 before we describe the calibration techniques in detail.

Table 2.4: Properties of a sensor that needs to be balanced when choosing a sensor.

Sensor Properties	**Description**	**Features of Expensive Sensors**	**Features of Low-Cost Sensors**
Specificity	Ability to measure only the gas of interest and not have a response to other gases	Very specific	Cross-sensitivity with other gases is a major concern
Sensitivity	Measuring the expected highest and lowest gas concentrations	Reasonable scaling of sensitivity	Unable to maintain noise-free measurements at one end while adjusting the highest or lowest points
Reliability	Property of the instrument to be used continuously	Highly reliable and usually requires less frequent maintenance visits	Less reliable
Reproducibility	Keeping the measured sample gas unaltered in chemical and physical composition after the measurement is taken	Usually reproducible	Show problems with reproducibility in many cases
Stability	Ability to generate accurate responses in longer unattended measurement periods	Can be calibrated and then left to operate unattended for a longer time	Errors get introduced in the measurements with passing time
Response time	The time for sampling over which measurement samples are taken	Usually lower (with exceptions)	Usually higher (with exceptions)
Accuracy and precision	Closeness of the measured values to the actual value and to each other	Highly accurate and very precise	Lower accuracy and precision
Cost	Cost of the sensor equipment, setup, energy consumption, and operation	Very expensive	Lower overall cost

2.3.2 Classification of Low-Cost Sensors

Researchers have reported several types of cost-effective gas sensing technologies in the last two decades [45-47, 104]. Some of the most commonly used methods for gas concentration measurements are often based on electrical variation with different materials and standard methods and techniques including metal oxide (MOX) semiconductor [95, 105-113], polymer [100-102], carbon nanotube [114, 115], moisture absorbing materials [116, 117] and electrochemical (EC) gas sensing [25, 49, 94, 100, 118-120] etc. Many methods based on other non-electrical variations have been also reported for gas sensing such as optical methods [121-126], acoustic methods [127-129], gas chromatography [130-132], calorimetric methods [133-135], etc.

The LCS technology can be roughly categorised into four different categories [104] as shown in Figure 2.2.

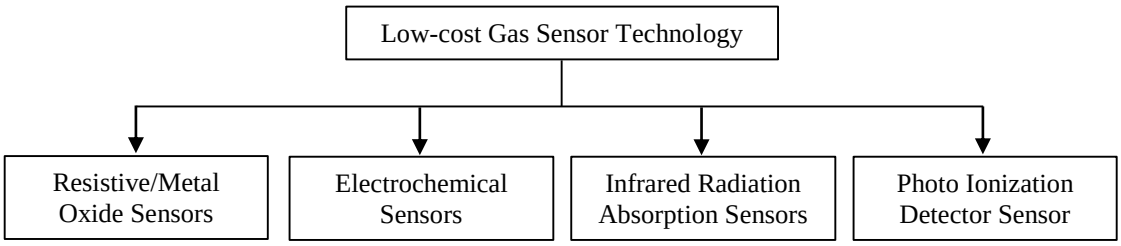


Figure 2.2: Classification of available LCS technologies.

Table 2.5 shows the various advantages and disadvantages of these four types of LCS technologies. Two of the most widely reported low-cost gas sensors to measure the ambient air pollution are Metal-oxide-semiconductor (MOS) or Metal Oxide (MOX) sensors [56, 79, 95, 113, 136-138] and Electrochemical (EC) gas sensors [9, 25, 28, 49, 94, 100, 118, 119]. Although, MOX sensors are light-weight (a few grams), small in size (few millimetres) and also have wide commercial availability at low-cost (~ $10) [42, 104], these sensors have a non-linear response concerning both the target and the interfering gases [139] and the sensitivity suffers from both long and short term drift [56, 137]. MOX sensor response is reported to become slower when recovering to the baseline [140] and shows difficulty with reproducibility [104]. Also such sensors suffer from cross-sensitivity to other (interfering) gases and changes in environmental conditions [9, 56]. Although MOX sensors are widely used as LCS to monitor air pollution, the EC

sensors, on the other hand, have been recognised as better suited for real-time ambient air quality monitoring in some research [9, 94, 120].

The EC sensors employ an electrochemical reactions with the target gas and a current is generated in the electrolyte that depends on the gas concentration [9, 141]. This current is the measurement for the concentration of gas as the EC sensor response is usually linear or logarithmic [104]. EC sensors also suffer from non-linear responses, cross-sensitivity to other (interfering) gases. EC sensors are about few tens of millimetres in size and costs around ~$100 [42, 104].

Table 2.5: Properties of different low-cost gas sensing technologies.

Low-Cost Gas Sensors	**Advantages**	**Disadvantages**
Metal Oxide Sensors	• Response to a wide range of concentrations (from ppb to ppm) [142, 143] • Small size (around a dozen millimetres) and weight is about a few grams [104] • Can operate at a higher temperature and have longer lifetime compared to other LCS with similar cost [144]	• Need high temperature to increase the rate of reactions [104] • Response time tends to be slow (around a few minutes) [143] • Highly cross-sensitive [79] • Stability is the least among all types of sensors [95] • Need to be recalibrated from time to time due to drift [95, 104] and require additional resources • Difficulties with reproducibility [79]
Electrochemical Sensors	• Small size (about 20 millimetres) [104] • Low power consumption [9, 144] • Low detection limit (LOD) of gas concentration [143] • Error tends to be similar in all sensors and show smaller deviation from linearity [118] • Show slightly better long term stability [28] • Sensor response is usually linear or logarithmic [25] • Cross-sensitivity can be fine-tuned by selection of different electrode materials [94]	• Ambient temperature and humidity have an influence on the sensor response [14, 49] • Affected by wind speed in the outdoor environment [28, 120] • High sensitivity [144] • Cross-sensitivity can cause error in response [14] • Signal drift can be an issue for long-term deployment [145]

Infrared Radiation Absorption Sensors	• Very small size (a few millimetres) [104] • Power consumption is low (~100mW) [104] • Very fast response time (~10s) [146] • Large detection range; especially for hydrocarbon gases and CO_2 [146] • Very good CO_2 sensor for having a very characteristic absorption band [147]	• Noise is high for sensing hydrocarbons [104, 146] • Cross-sensitive for hydrocarbon gases as they share similar absorption band [104]
Photo Ionization Detector Sensors	• Power consumption in low (~10mW) [104] • Very good sensors for volatile organic compounds (VOCs) and Benzene gas [46] • Low deviation from the linear response for lower gas concentrations [46] • Fast response time (~3s) [104]	• Slightly larger and heavier instrument compared to other sensor types [104] • Temperature and humidity affect the sensor response [46] • Require very frequent calibration [104]

Both MOX and EC sensors inherently suffer from low selectivity and poor specificity [99]. As a result, these sensors are cross-sensitive to different interfering components that exist in the air and thus do not exclusively measure their target gases [20]. This is exacerbated while measuring gaseous pollutants (e.g., O_3, NO_2, CO, SO_2) in outdoor environment [95, 97]. Such errors in LCS are mainly caused by various physical properties of the sensor and pure chemical interferences, and require sophisticated calibration approaches [21].

As mentioned earlier, the common drawback of these LCS is that their responses change with changing environmental conditions [36, 148]. Reports from different laboratory and field tests suggest that LCS performance is significantly worse in real-world deployment compared with controlled lab environments [21]. Temperature and relative humidity can change the sensor measurements significantly [28, 120]. Dependency on ambient conditions such as temperature, causes non-linear responses in LCS [28] as well as any large changes in relative humidity causes change in the sensitivity of commonly used MOX and EC sensors [51, 95, 119].

2.3.3 Common Errors in LCS

One of the main reasons of LCS not being able to replace costly equipment and instruments (reference sensors) for air quality monitoring is comparatively lower accuracy [11, 18]. Although it is not expected from LCS to generate the same results of an expensive device while measuring air pollutant concentrations, it should exhibit a strong correlation.

The effects of ambient conditions and interference from other pollutants cause a combination of different types of errors in LCS, namely, dynamic boundaries [9, 21], systematic errors [11, 18], non-linear response [149, 150] etc. These errors are usually related to the working principle of the sensor, used hardware, measurement technique etc. [21]. A brief description of these errors is given in Table 2.6.

According to many reports, systematic errors and signal drift are the two most commonly found errors in commercially available sensors for ambient air quality monitoring [99, 103]. EC and MOX gas sensors generally exhibit dynamic boundary errors and non-linear response respectively [79, 120]. Ambient conditions and cross-sensitivity can directly cause these errors. However, pin pointing an error to a particular reason especially for LCS is very hard as the errors are the product of multiple inherent properties of LCS.

A commonly used approach to minimize the errors or the combination of errors in these pollution monitoring LCS is calibration. Several different calibrations methods have been reported by researchers to reduce the effect of sensor errors which are discussed in the next section.

Table 2.6: Different types of errors in LCS.

Errors	Source of Error	Effects	Major Concern
Dynamic Boundary	Range of a pollutant concentration in which a sensor is sensitive to; e.g., limit of detection (LOD) [151]	Getting affected by high noises [21]	LCS are significantly affected by low SNR (signal to noise ratio) at low concentrations [9, 49]
Systematic Error	Due to imperfect calibration parameters and are usually not related to the sensing principle [21]	A constant offset over whole concentration range or an over/under-estimation of concentration in certain range [11, 30, 31]	Factory calibrated sensors [21]
Non-linear Response	Caused by dependency on environmental conditions and internal signal processing of the sensor [28]	Over/under-estimation of sensor responses with increasing or decreasing sensor outputs [17]	LCS are well known for delivering a non-linear response [149, 150, 152]
Signal Drift	Aging and impurity affecting the sensor performance over a long time [153]	Causes a slow drift of the sensor sensitivity and as a result the sensor measurements changes over time [96]	Most common and hampering error with low-cost sensors in long-term deployments [154]

2.4 Sensor Calibration

As discussed in the previous section, the data provided by LCS have accuracy issues while measuring ambient air quality. Researchers are working to develop appropriate sensor calibration methods in order to improve the accuracy of existing and future air quality monitoring deployments [21]. Through calibration, measurements of LCS are transformed in a way so that the calibrated measurements closely agree with the "reference" measurements from a reliable high quality sensor. Although sensor calibration work has been going on for many years, recent developments in the low-cost air quality monitoring sensors along with the introduction of using Internet of Things (IoT) protocols in environmental monitoring [90, 155, 156] has significantly intensified the interest of researchers. Sensor calibration can potentially play a huge role in the deployment of the LCS by ensuring that the gap between the results from LCS and an

expensive high quality sensor is reduced down to tolerable level [34, 51, 52]. Calibration models are the fundamental keys to overcome the major shortcomings of low-cost sensors. In the following sections, calibration setups, models and strategies for low-cost air pollution sensor deployment found in literature are reported.

2.4.1 Calibration Setup

Calibration models for air quality LCS are strategies followed to compare its raw measurement readings with that from a reliable reference sensor and transforming them to more accurate calibrated values. These calibration models vary depending on the type and number of sensors used in the measurement [15], source of the error [50], placement and lifetime of the sensor [37], and availability of the reference sensor or setup with accurate measurements [68]. The calibration setup can be classified into two general groups (Figure 2.3): laboratory setup under a controlled environment and field setup under an observed environment (sensor co-located at reference monitoring site) [21, 51, 153].

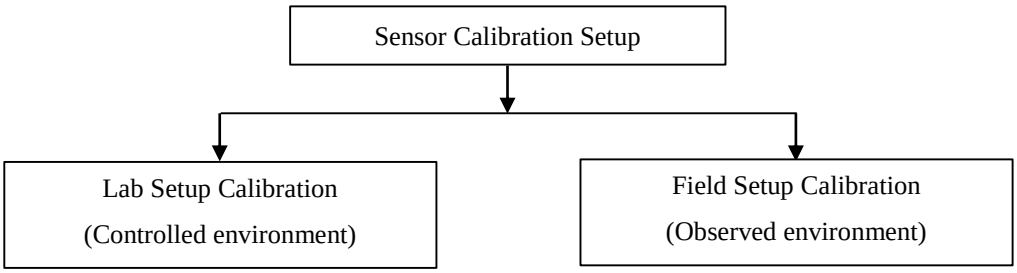


Figure 2.3: Classification of calibration models on the basis of setup environment.

Although lab setups are faster and convenient, many researchers recommend field setups for calibration [37, 42, 99, 150, 157]. While lab based calibration allows stringent control over the concentration ratios and ranges of pollutants [51, 152, 158], field calibration takes into account the inherently complex nature of the ambient conditions [34, 45, 50, 148]. In order to properly characterize sensor responses to which the sensors are subjected to, the calibration approach should consider the number of pollutants present in the surrounding and their interference concentrations [50, 94, 100]. It is difficult to emulate the actual ambient conditions in the lab setup because there are too

many parameters to control [51, 119, 152]. On the other hand, sensors are exposed to situations with realistic environmental conditions during field testing setups [30, 31]. Calibration that involves more than one ambient factor with on-field data are being currently reported as an efficient tool for improved performance regarding real-time air quality monitoring [17, 36, 50, 57, 159].

2.4.2 Different Types of Sensor Calibration

Researchers have been working to develop calibration techniques to improve the accuracy of LCS. As discussed, such sensors can be calibrated by co-locating them with accurate sensors so that the calibrated measurements of the low-cost sensor closely agree with the co-located accurate reference sensor [21] and the co-located measurements are often performed during "field deployment". The key component of the calibration is the training of regression models to capture the complex, often nonlinear, relationship between the raw sensor output and the ground truth provided by the accurate reference sensor. Three types of corrections have been widely reported in the literature for calibrating LCS measuring air pollutants:

- Offset and Gain Correction
- Temperature and Humidity Correction
- Cross Sensitivity Correction

2.4.2.1 Offset and Gain Correction

Offset and gain correction maps the raw measurements of sensing to a target pollutant concentration [21]. It also deals with the errors associated with dynamic boundaries, systematic errors and some potential non-linear responses [21, 160]. Calibration models for the offset and gain correction can be derived in both field and lab setups [161] by co-locating with a (reliable) reference sensor that provides actual measurement of the target pollutant concentration (ground truth). Offset and gain correction is a univariate regression problem so that the calibrated gas sensor output is:

$$X_{calibrated}^{OG} = \Phi^{OG}\{\mathrm{X}_{raw}\} \quad 2.3$$

Φ^{OG}, is derived solely based on the raw gas sensor input (X_{raw}) from the LCS to minimize the difference between calibrated sensor output $X^{OG}_{calibrated}$ and the ground truth provided by the co-located reference sensor.

Commercially available LCS typically come with an initial laboratory calibration for offset and gain correction. Manufacturers usually provide the sensor responses over a range of target pollutant concentrations or calibration curves recorded in a lab setup [21]. Researchers have favoured linear regression methods for linearly correlated sensor data [21] and non-linear curve fitting for exponential or power law gain terms [149, 152].

2.4.2.2 Temperature and Humidity Correction

Sensor response or sensitivity of LCS can be affected by changes in the ambient temperature and relative humidity (RH) [120]. EC sensors are typically configured so that the reactions are diffusion-limited, and the diffusion coefficient is affected by temperature [68, 162]. Similarly, relative humidity (RH) also affect the measurements, especially with higher values (RH>75%), significant error can be observed [120, 150]. Long term observation using EC sensors have revealed that changes in temperature affects the sensor response [28]. Literature also reports sensitivity losses due to the change in ambient temperature [9]. Similar loss of sensitivity was experienced under changing relative humidity (RH). For example, Wang *et al.* observed the effect of relative humidity on sensitivity of a MOX sensor and reported significant reduction in sensitivity when placed at extreme relative humidity of 95% from dry air conditions [51]. Likewise, Pang *et al.* observed a noticeable drop in sensitivity (~20%) when the relative humidity was increased to 85% from 15% [119]. Therefore it is highly recommended to perform temperature and humidity correction for low-cost sensors [120].

The process of correcting temperature and humidity is facilitated by the availability of inexpensive but reliable temperature and humidity sensors on board the gas sensor unit. Many calibration approaches leverage these additional measurements and convert the single-variate models of offset and gain calibration to multi-variate models so that Equation 2.3 (from offset & gain) can be updated to following Equation 2.4:

$$X^{OG,T/H}_{calibrated} = \Phi^{OG,T/H}\{X_{raw}, T, H\} \quad 2.4$$

The regressor $\Phi^{OG,T/H}$ is now derived from three input variables, raw sensor data, temperature (T) and humidity (H), to minimize the difference between the calibrated output $X_{calibrated}^{OG,T/H}$ and the ground truth.

Multiple linear regression has been used in the past to approximate a linear combination of raw air pollutant concentration, temperature and humidity measurements that best fits the target reference concentration for low-cost sensors [15, 18, 42, 49, 98, 103, 148, 163]. Some other strategies include non-linear curve fitting [138, 150], in which a detailed model of the physical effects of ambient temperature is derived for the sensors. Popoola *et al.* reported a correction algorithm for temperature baseline in EC sensors [28]. They observed that, while CO had a linear relationship to its reference, the NO sensor showed a strong exponential relationship. Therefore their proposed model used a linear line fitting for the CO sensor and an exponential curve fitting for the NO sensor to correct temperature effects on the measurement [28]. Similar exponential curve fitting for temperature and humidity correction has been reported in other studies [138, 164]. However, modern machine learning techniques have started to become popular for correcting temperature and humidity effects on the LCS [68].

2.4.2.3 Cross Sensitivity Correction

Low-cost gas sensors are particularly sensitive towards other gases in the atmosphere and their responses are required to be corrected for this cross-sensitivity [21]. To do so, a sensor node (often termed sensor array) with multiple gas sensors measuring multiple pollutant concentrations is assembled [17, 18, 21, 37, 49, 68]. The calibration model can now be written by updating the previous Equation 2.4 as:

$$X_{calibrated}^{OG,T/H,CS} = \Phi^{OG,T/H,CS}\{\mathrm{X}_{raw}, T, H, \mathbf{X}_{cross}\} \quad 2.5$$

In Equation 2.5, $\mathbf{X}_{cross}$ represents other pollutant data measured by the sensor array alongside the raw target pollutant measurement X_{raw}, T and H that are used to derive the regression model $\Phi^{OG,T/H,CS}$ that produces the calibrated output $X_{calibrated}^{OG,T/H,CS}$.

Popular sensor array calibration methods are linear regressions [30, 31, 37, 99, 119] and machine learning algorithms [17, 57], a list of which can be found in Table 2.7.

One of the most common examples of cross-sensitivity in EC sensors is between NO_x and O_3 [30, 36, 97, 99, 119]. Pang *et al.* reported the compensation for the influences of NO_x concentrations on ambient O_3 measurements in EC sensors by applying linear regression [119]. Similar approach was also performed by Maag *et al.* where O_3 concentrations were compensated for NO_2 measurements in an EC sensor [99]. Similarly, the cross-sensitivity errors of NO_2 sensor were reduced in low-cost MOX sensors by incorporating measurements of an additional O_3 sensor in the calibration process [21].

Sensor array calibrations facilitated by field co-location have provided the best performance regarding stability, accuracy and precision [15, 17, 68, 94]. However, this requires a large-scale analysis of concurrent data from multiple sensors [15, 21]. In addition, cross-sensitivity becomes more complex as different factors in the environment also affect the sensor responses [15, 45, 160]. More sophisticated training models are needed to manage this complexity in order to calibrate the LCS for real-world deployments. Thus, different machine learning methods are becoming more and more popular for calibration to compensate for the errors caused by cross-sensitivity, temperature and humidity effects present in LCS. In spite of recent developments in the calibration of air quality monitoring sensor arrays, selecting the optimal sensor array calibration to tackle all types of errors in outdoor environment remains a challenge. Table 2.7 summarizes different calibration setups and methods used by research groups for various sensor types and pollutants.

Table 2.7: Different calibration techniques used by various research groups as reported in the literature.

Pollutant	**Setup**	**Sensor Type**	**Method**	**Reference**
CO, NO_X	Field	MOX	Linear Regression (LR/MLR)	Carotta *et al.* 2001 [157]
NO_2	Field and Lab	MOX	Linear Regression (LR/MLR)	Tsujita *et al.* 2005 [138]
O_3, NO_2	Lab	MOX	Neural Networks (MLP)	Kamionka *et al.* 2006 [165]
C_6H_6	Field	MOX	Neural Networks (MLP)	De Vito *et al.* 2008 [48]
CO, NO_2, NO_X	Field	MOX	Neural Networks (MLP)	De Vito *et al.* 2009 [36]

CH_4	Field	MOX	Linear Regression (LR/MLR)	Eugster *et al.* 2012 [163]
O_3	Lab	MOX	Linear Regression (LR/MLR)	Hasenfratz *et al.* 2012 [166]
C_2H_5OH	Field	MOX	Gaussian Processes Regressors (GPR)	Monroy *et al.* 2012 [136]
CO, NO_2, NO_X	Field	MOX	Neural Networks (MLP)	De Vito *et al.* 2012 [167]
NO, NO_2, CO, SO_2,O_3	Field	EC, MOX	Support Vector Regression (SVR)	Nieto *et al.* 2013 [168]
CO, NO, NO_2	Field and Lab	EC	Linear Regression (LR/MLR)	Mead *et al.* 2013 [25]
C_2H_5OH	Field	MOX	Gaussian Processes Regressors (GPR)	Monroy *et al.* 2013 [137]
CO, NO_2	Field and Lab	MOX	Linear Regression (LR/MLR)	Piedrahita *et al.* 2014 [42]
C_2H_5OH, C_2H_4, C_3H_6O	Lab	MOX	Recurrent Neural Networks (RNN-RC)	Sheik *et al.* 2014 [169]
CO_2, CO, O_3, NO_2	Field and Lab	MOX	Linear Regression (LR/MLR)	Piedrahita *et al.* 2014 [42]
CO, C_2H_4	Lab	MOX	Recurrent Neural Networks (RNN-RC)	Fonollosa *et al.* 2015 [76]
O_3, CO	Field	MOX, EC	Linear Regression (LR/MLR)	Saukh *et al.* 2015 [170]
O_3, NO_2	Field	MOX	Linear Regression (LR/MLR)	Lin *et al.* 2015 [171]
O_3, NO_2	Field	EC	Linear Regression (LR/MLR)	Spinelle *et al.* 2015 [55]
CO	Field and Lab	MOX	Linear Regression (LR/MLR)	Masson *et al.* 2015 [150]
CO, NO, NO_2, O_3	Field and Lab	EC	Linear Regression (LR/MLR)	Gerboles *et al.* 2015 [172]
O_3, NO_2	Field	MOX, EC	Linear Regression (LR/MLR), Neural Networks (MLP)	Spinelle *et al.* 2015 [30]
NO_2, O_3, CO	Field	MOX, EC	Linear Regression (LR/MLR)	Maag *et al.* 2016 [99]
CO, NO, NO_2	Field and Lab	EC	Linear Regression (LR/MLR)	Popoola *et al.* 2016 [28]

CO, SO_2, NO, O_3, NO_2, CO_2, H_2	Field and Lab	EC	Gaussian Processes Regressors (GPR)	Lewis *et al.* 2016 [100]
O_3	Lab	MOX	Linear Regression (LR/MLR)	Spinelle *et al.* 2016 [56]
O_3, CO, NO_2, SO_2	Field	EC, MOX	Neural Networks (MLP), Random Forest Regression (RFR)	Borrego *et al.* 2016 [34]
NO, NO2, O3, CO	Field	EC	Neural Networks (MLP)	Esposito *et al.* 2016 [50]
CO, NO2, O3	Field and Lab	EC	Linear Regression (LR/MLR)	Sun *et al.* 2016 [33]
O_3, NO_2, NO, CO, SO_2	Field	EC, MOX	Linear Regression (LR/MLR)	Jiao *et al.* 2016 [18]
CO, NO, O_3, NO_2	Field and Lab	EC	Linear Regression (LR/MLR)	Castell *et al.* 2017 [11]
CO, NO, NO_2, O_3	Field	EC	Linear Regression (LR/MLR), High-Dimensional Model Representation (HDMR)	Cross *et al.* 2017 [94]
NO2	Field	EC	Linear Regression (LR/MLR)	Mijling *et al.* 2017 [173]
NO_2	Field	MOX	Linear Regression (LR/MLR)	Fang *et al.* 2017 [37]
O_3, NO_2	Field and Lab	EC	Linear Regression (LR/MLR)	Sun *et al.* 2017 [103]
NO, CO, CO_2	Field	MOX, EC	Linear Regression (LR/MLR), Neural Networks (MLP)	Spinelle *et al.* 2017 [31]
O_3	Field and Lab	EC	Linear Regression (LR/MLR)	Pang *et al.* 2017 [119]
NO, NO_2, CO	Field	MOX	Neural Networks (MLP)	Barakeh *et al.* 2017 [174]
O_3, NO_2	Field	EC	Linear Regression (LR/MLR)	Mueller *et al.* 2017 [96]
CO, NO, NO_2, NO_X, O_3, SO_2	Field	EC	Linear Regression (LR/MLR), Neural Networks (MLP), Gaussian Processes Regressors (GPR), Support Vector Regressor (SVR), Recurrent Neural Networks (RNN-RC)	Esposito *et al.* 2017 [58]
CO, NO, O_3, NO_2	Field and Lab	EC	Linear Regression (LR/MLR)	Wei *et al.* 2018 [120]

O_3, CO_2	Field and Lab	MOX, EC	Linear Regression (LR/MLR), Neural Networks (MLP)	Maag *et al.* 2018 [175]
O_3, CO, NO_2, SO_2	Field	EC, MOX	Neural Networks (MLP), Random Forest Regression (RFR)	Borrego *et al.* 2018 [26]
O_3	Field	MOX	Linear Regression (LR/MLR)	Barcelo-Ordinas *et al.* 2018 [15]
SO_2	Field	EC	Linear Regression (LR/MLR)	Hagan *et al.* 2018 [49]
CO, NO_2	Field	EC	Linear Regression (LR/MLR), Gaussian Processes, Neural Networks (MLP), Support Vector Regressor (SVR), Recurrent Neural Networks (RNN-RC)	Esposito *et al.* 2018 [57]
NO_2	Field	EC	Linear Regression (LR/MLR), Radom Forest, Support Vector Regressor (SVR), Neural Networks (MLP)	Cordero *et al.* 2018 [65]
CO, NO_2, CO_2, O_3	Lab and Field	EC	Linear Regression (LR/MLR), Random Forest Regression (RFR)	Zimmerman *et al.* 2018 [68]
NO_2, CO	Field	MOX, EC	Linear Regression (LR/MLR), Gaussian Processes Regressors, Support Vector Regressors (SVR), Neural Networks (MLP), Recurrent Neural Networks (RNN-RC)	De Vito *et al.* 2018 [17]
SO_2	Field	EC	Linear Regression (LR/MLR), K-Nearest Neighbors (KNN)	Hagan *et al.* 2018 [49]
NO, NO_2	Field	EC	Linear Regression (LR/MLR), Support Vector Regressors (SVR), Random Forest Regression (RFR)	Bigi *et al.* 2018 [67]
CO, O_3	Field	EC	Linear Regression (LR/MLR), Neural Networks (MLP)	Topalović *et al.* 2019 [53]
CO, NO, NO_2	Field	EC	Linear Regression (LR/MLR)	Munir *et al.* 2019 [176]
CO, NO, NO_2, O_3	Field	EC	Linear Regression (LR/MLR), Gaussian process Regression, Neural Networks (MLP), Random Forest Regression (RFR)	Malings *et al.* 2019 [71]
CO, NO, NO_2, O_3,	Field	EC	Linear Regression (LR/MLR)	Karagulian *et al.* 2020 [177]

CO	Lab	MOX	Neural Networks (MLP), Support Vector Regressors (SVR), Recurrent Neural Networks (RNN-GRU)	Wang *et al.* 2020 [178]
CO, NO_2, O_3	Field	EC	Neural Networks (MLP)	De Vito *et al.* 2020 [16]
CO, NO_2, O_3	Lab	MOX	Support Vector Regressors (SVR)	Djedidi *et al.* 2021 [66]
CO, NO_2	Lab and Field	EC	Linear Regression (LR/MLR), Neural Networks (MLP)	De Vito *et al.* 2021 [10]
NO_2	Field	EC	Convolution Neural Network (CNN), Linear Regression	Vajs *et al.* 2021 [179]
CO, NO, NO_2, O_3	Field	EC	Linear Regression (LR/MLR)	Zuidema *et al.* 2021 [61]
CO, NO_2, O_3	Field	EC	Linear Regression (LR/MLR)	Liang *et al.* 2021 [14]
NO_2	Field	EC	Linear Regression (LR/MLR)	Laref *et al.* 2021 [145]
NO, NO_2, PM10	Field	EC	Linear Regression (LR/MLR)	Wahlborg *et al.* 2021 [180]
CO, NO_2, SO_2, O_3	Field	EC	Neural Networks (RNN-LSTM)	Han *et al.* 2021 [77]
O_3	Field	EC	Random Forest Regression (RFR), Gradient Boosting Regression (GBR), Extreme Gradient Boosting (EGB)	Bagkis *et al.* 2022 [69]
NO_2	Field	EC	Linear Regression (LR/MLR)	Rogulski *et al.* 2022 [60]
O_X, NO, NO_2, CO	Field	EC	K-Nearest Neighbors (KNN), Random Forest Regression (RFR), High-Dimensional Model Representation (HDMR), Quadratic Regression (QR), Linear Regression (LR/MLR)	Bittner *et al.* 2022 [70]
NO_2	Field	EC	Linear Regression (LR/MLR)	Hofman *et al.* 2022 [19]
O_3	Field	EC, MOX	Linear Regression (LR/MLR)	Badura *et al.* 2022 [59]
CO	Lab and Field	EC	Bayesian Linear Regression (BLR), Neural Networks (MLP)	Tancev *et al.* 2022 [181]

CO, NO_2, SO_2, O_3	Field	EC	Linear Regression (LR/MLR), Random Forest Regression (RFR), Gaussian Process Regression (GPR)	Martín-Baos *et al.* 2022 [73]
CO, NO_2, SO_2, O_3	Field	EC	Linear Regression (LR/MLR)	Daepp *et al.* 2022 [182]
CO, CO_2, NO_2, SO_2	Field	EC	Linear Regression (LR/MLR)	Jabbar *et al.* 2022 [183]
NO_2, O_3	Field	EC	Neural Networks (MLP)	Ganji *et al.* 2023 [184]
NO_2	Field	EC	Linear Regression (LR/MLR), Random Forest Regression (RFR)	Wang *et al.* 2023 [72]
O_3	Field	EC	Linear Regression (LR/MLR), Neural Networks (MLP), Gradient Boosting Regression (GBR), Adaptive Boosting (AB), Extreme Gradient Boosting (EGB)	Sa *et al.* 2023 [74]
SO_2	Lab and Field	EC	Linear Regression (LR/MLR)	Rivero *et al.* 2023 [62]
CO_2,	Field	EC	Linear Regression (LR/MLR)	Kappelt *et al.* 2023 [64]
CO, NO_2, O_3	Field	EC	Linear Regression (LR/MLR), Polynomial Regression (PR), Random Forest Regression (RFR)	Hasan *et al.* 2023 [185]

The pie chart in Figure 2.4 summarizes the information presented in Table 2.7 where different calibration methods are shown with their overall percentage of use. It can be observed that almost half of the used calibration methods in the literature are linear regressions (or multiple linear regressions). MLP has been also frequently used in the previous works, just over 17%. RFR, SVR and RNN (including RC, GRU and LSTM) are the only other calibration methods that have been utilized over 5%. All other mentioned algorithms (e.g., GBR and CNN) have been utilized only on a few occasions. The literature review shows that among the NN-based techniques, CNN or more specifically, One Dimensional Convolutional Neural Network (1DCNN) has not been well investigated for LCS calibration. 1DCNN has demonstrated excellent performance for a variety of applications (e.g., indoor localisation [186], human activity recognition [187], time series forecasting [188]). However, there are only two reports [179, 189] of

1DCNN being utilized for the calibration of air pollution monitoring. Kureshi *et al.* [189] employed it for the calibration of Particulate Matter (PM) sensors. In a recent publication that investigated the impact of the pandemic on the air quality, Vajs *et al.* [179] employed 1DCNN to calibrate low-cost NO_2 and PM sensors. However, they did not benchmark its performance against any other ML techniques; therefore, it is difficult to ascertain its (comparative) efficacy. Similarly, GBR, an ensemble learning technique, has also not been widely utilized for ambient gas pollutant sensor calibration although it has shown good performance in other applications (e.g., PM sensor calibration [190], and prediction and forecasting [191, 192]). While Bagkis *et al.* [69] and Sa *et al.* [74] employed GBR for gas sensor calibration, their works mainly focused on temporal drift correction and the performance of GBR was not benchmarked against sophisticated techniques like NNs.

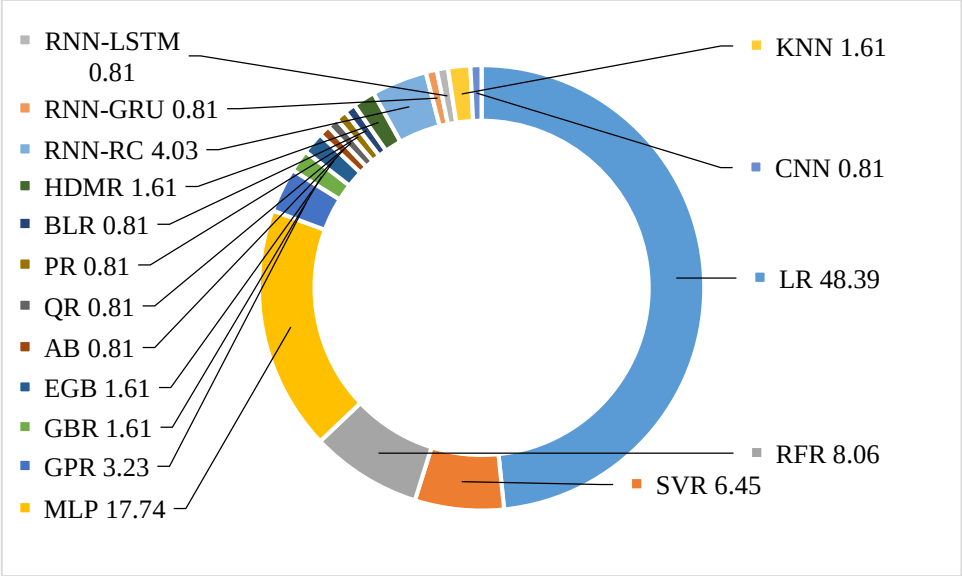


Figure 2.4: Summary of used calibration methods found in the literature. The numbers beside the calibration methods show the percentage of being utilized overall.

CHAPTER 3

METHODOLOGY

This chapter describes the methodology employed for this research. The various aspects of the gas pollutant datasets that were utilized for the study are described in this chapter. This is followed by a detailed, step by step description of the calibration process. The proposed calibration algorithms and performance metrics that are used to evaluate the efficacy of the various algorithms are also described in this chapter.

3.1 Dataset Utilized

For this study, three different datasets were used to develop and test novel machine learning based calibration algorithms. The datasets were acquired from published online open sources or via a direct request to the corresponding research groups. Table 3.1 provides an initial summary of the datasets. Figure 3.1 shows an approximate map of the deployment of these LCS across the globe.

Table 3.1: Summary of the datasets.

Dataset	Target Pollutant	LCS array	Collection Source	Location	Reference
Dataset 1	CO, NO_2	MOX	Online open source	Lombardy Region, Italy	[17, 36, 48, 167]
Dataset 2	CO, NO_2	EC	Acquired through request	Naples, Italy	[10, 16]
Dataset 3	CO, NO_2	EC	Acquired through request	Guangzhou, China	[14]

Figure 3.1: Locations of the LCS used in this study.

3.2 Dataset Description

The focus was on the calibration of CO and NO_2 sensors. As discussed in Chapter 2, these gases are important components of the Air Quality Index (AQI) [156]. Both the reference (ground truth) data from a co-located accurate sensor and raw (unchanged electrode data from LCS) data were available for all three deployments (Dataset 1, 2 and 3) [10, 14, 16, 36, 48, 167] for both pollutants. These deployments also collected other pollutant data using LCS. This will allow us to address cross-sensitivity. Temperature and relative humidity data are available for all three setups, allowing us to mitigate the temperature and relative humidity effects. A more detailed description of the datasets can be found in Table 3.2. Also, Figures 3.2, 3.3 and 3.4 illustrate the distribution (histogram) of the collected reference and LCS data along with box and whisker plots of temperature and relative humidity. These tables and figures clearly show the diversity of the utilized datasets and different of scales between the ground truth and the LCS responses.

Table 3.2: Details of the datasets.

Dataset	**Time Span (Days)**	**Location**	**LCS array**	**pollutants measured**	**Reference Sensor**
1 [17]	391	Lombardy Region, Italy	MOX	CO, NO_2, O_3, NMHC, NO_X (only working electrode [we] data)	Air pollution analyzer, operated by the regional environmental protection agency (ARPA)
2 [16]	965	Naples, Italy	EC	CO, NO_2, O_3 (both working [we] & auxiliary [ae] electrode data)	Teledyne™ 300 CO analyzer and Teledyne™ T200 NO_2 chemiluminescence analyzer
3 [14]	152	Guangzhou, China	EC	CO, NO_2, O_3 (both working [we] & auxiliary [ae] electrode data)	CO data were collected by a gas analyzer based on infrared absorption (Model 48i-TLE, Thermo Scientific, Waltham, MA, USA). NO_2 was measured by a chemiluminescence analyzer (Model 42i-TL, Thermo Scientific, Waltham, MA, USA)

3.2.1 Dataset 1

The dataset was recorded by a multi-sensor device [17] containing an array of five low-cost MOX sensors that measure CO, NO_2, O_3, Non-methanic Hydrocarbons (NMHC), and NO_X along with temperature (T) and relative humidity (RH). It includes 9357 samples of hourly averaged responses recorded between March 10, 2004, to April 04, 2005, Lombardy Region, Italy.

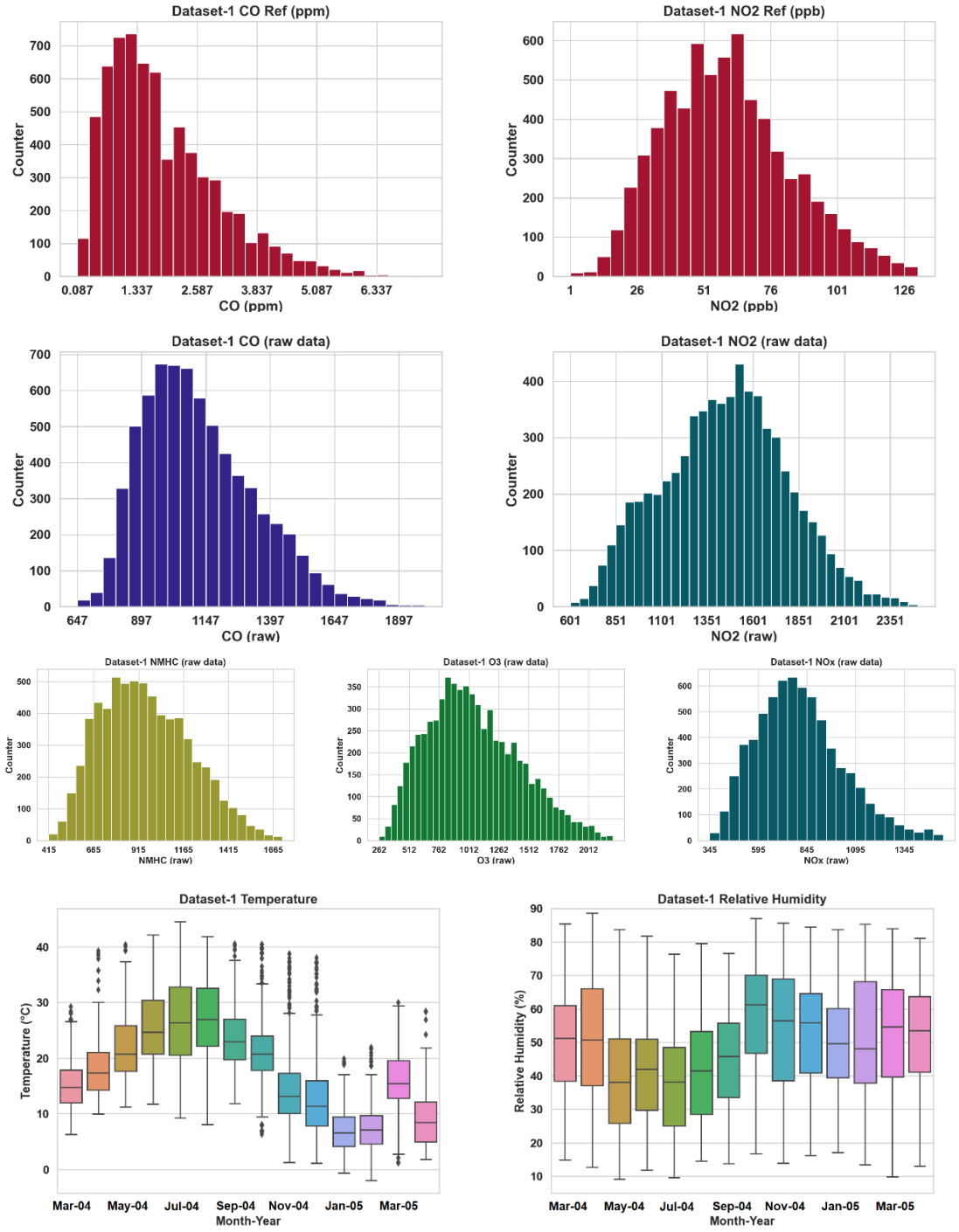


Figure 3.2: Distribution (histogram) of the reference and raw LCS data, box plots of temperature and relative humidity from dataset 1.

3.2.2 Dataset 2

This dataset includes the responses of a MONICA multi-sensor device [16] deployed in the Italian city of Naples. The gas sensor hardware consists of an array of electrochemical gas sensors to measure CO, NO_2, O_3, and T and RH. The 31 months' deployment collected a total of 13595 samples of hourly average from April 2018 to November 2020.

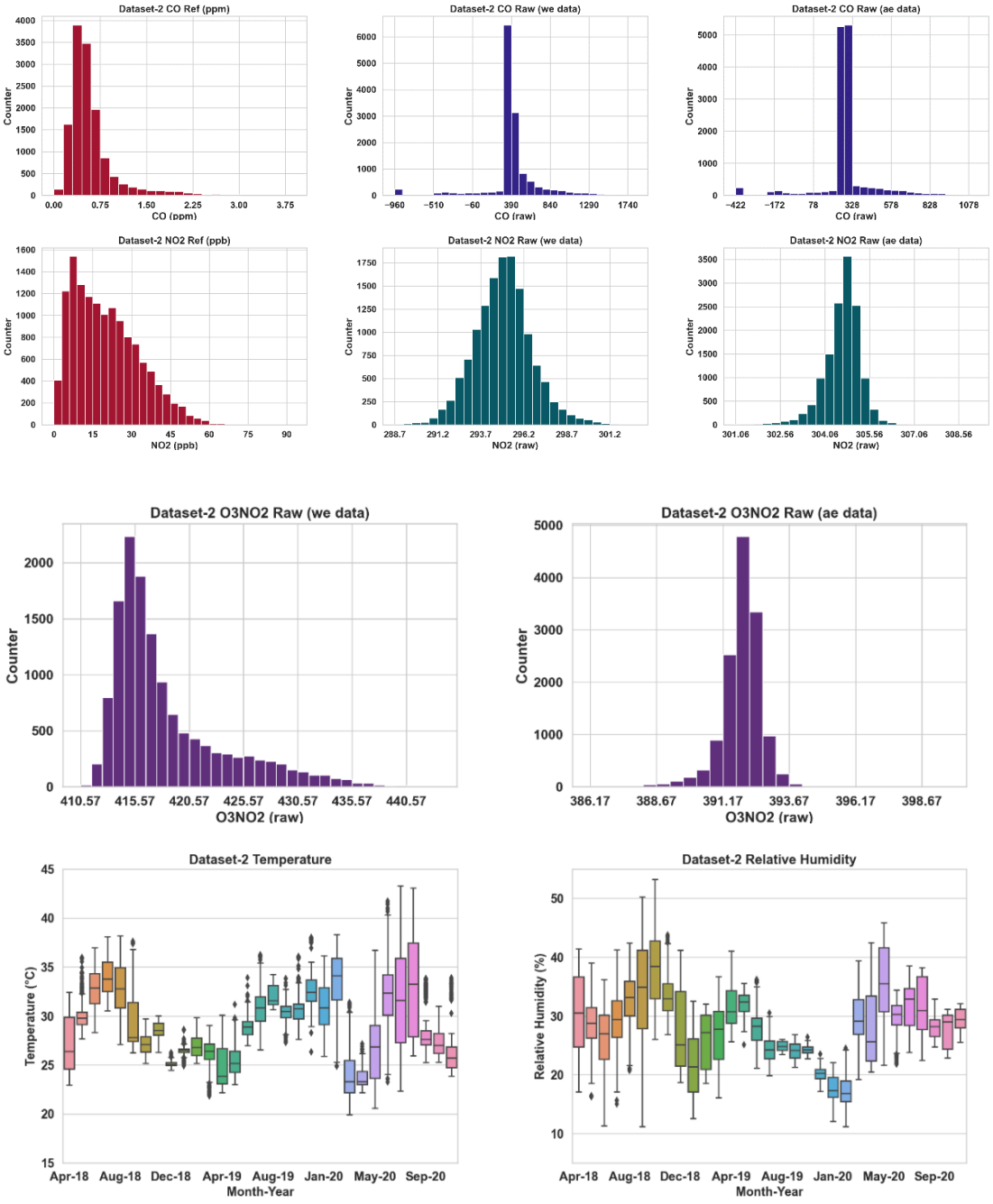


Figure 3.3: Distribution (histogram) of the reference and raw LCS data, box plots of temperature and relative humidity from dataset 2.

3.2.3 Dataset 3

This dataset was recorded by a multi-sensor device [14] deployed in the Chinese city of Guangzhou. The array of EC gas sensors measures CO, NO_2 and O_3 along with T and RH. Total of 4368 samples of hourly average data were collected over a span of six months between October 1, 2018, and March 1, 2019. More details of the dataset can be found in [14]. Please note that this dataset is also available at a higher per-minute, sampling rate.

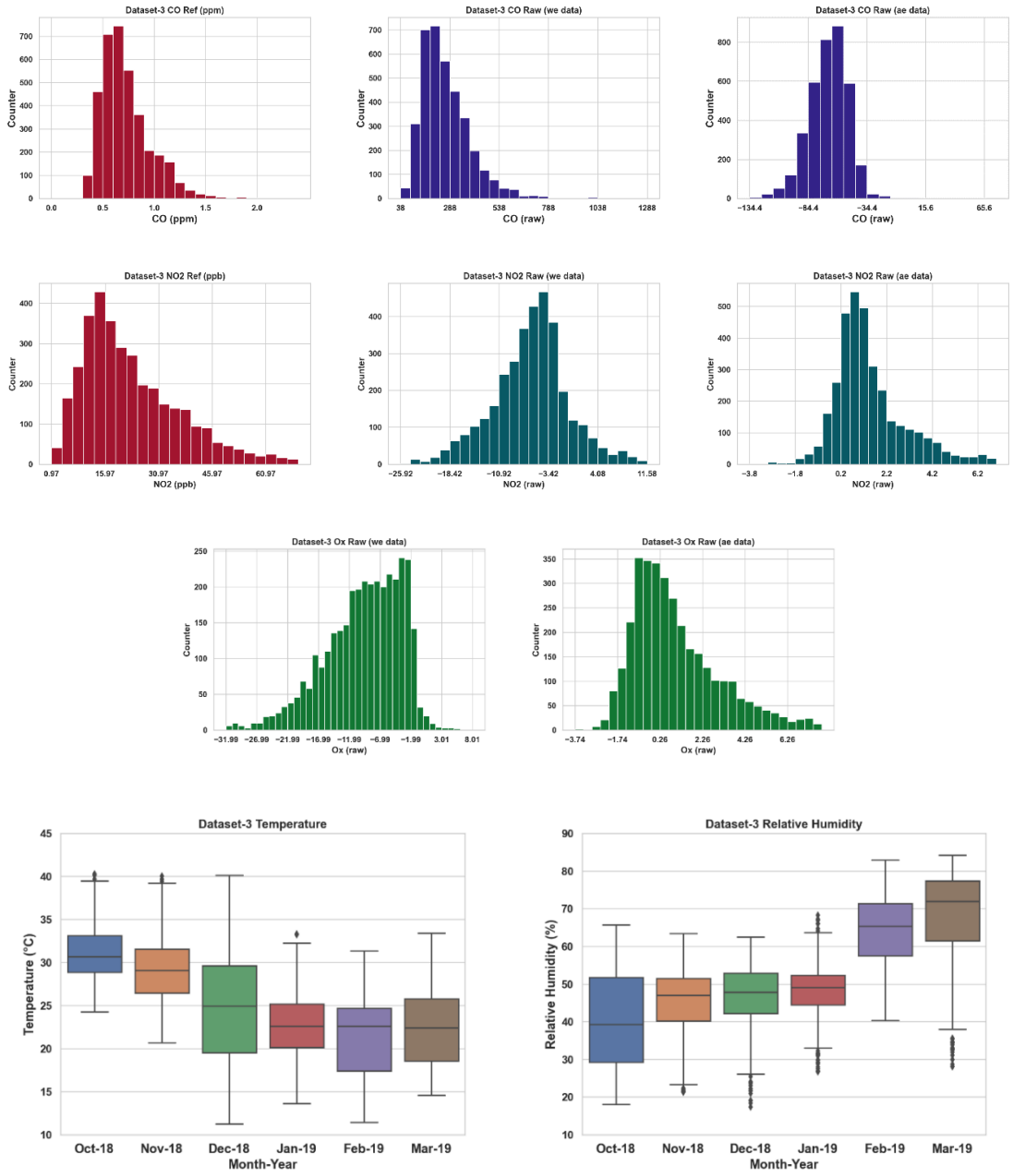


Figure 3.4: Distribution (histogram) of the reference and raw LCS data, box plots of temperature and relative humidity from dataset 3.

3.3 Dataset Cleaning

It was found that the data collected by the LCS and reference sensors have missing samples. Also, as discussed, having other pollutant data and temperature and relative humidity data are important for addressing the dependency to these co-variates in the gas measurements. Therefore, for any given instant, all pollutant (and temperature and relative humidity) data need to be available from the cost-effective sensor alongside the reference data for multivariate calibration. As a result, readings of selected time instants were removed from each dataset if any pollutant data (and temperature and relative humidity data) from the cost-effective sensors or the ground truth data were missing.

These datasets contained some extreme values in their samples which are outside the expected range or unlike the other samples. Performance can be improved in machine learning modelling by removing these outliers. The standard deviation of the sample can be used as a cut-off for identifying outliers if the distribution of values is a Gaussian or Gaussian-like distribution. Three standard deviations from the mean is a common cut-off in practice for removing outliers from the samples [193, 194]. In this case, the mean and standard deviation of the samples are calculated, and samples having more than three standard deviations from the mean are identified as outliers and removed from the dataset. The final sample numbers after data cleaning are given in Table 3.3.

Table 3.3: Number of samples used in the datasets after data cleaning.

Dataset	Target Pollutant	Number of samples
Dataset 1	CO	6941
	NO_2	6743
Dataset 2	CO	12982
	NO_2	12123
Dataset 3	CO	3639
	NO_2	3412

3.4 Feature Scaling

Feature scaling is the process of changing all the features in dataset into a similar range of values. It is one of the most important data pre-processing stages in machine learning as it helps algorithms to train and converge faster. A single dataset may have multiple features that may range in different values, and this can affect the algorithm training in unexpected ways. Hyperparameters or weights in the trained model can be particularly biased towards larger values. Thus, feature scaling is applied before the model training by fitting the scaler on the training data and then again using it to transform the testing data back to the original values.

The most commonly used feature scaling techniques are Normalization (or Min-Max Scaling) and Standardization (or Z-Score Normalization) [195]. However, there is no hard and fast rule that says when to apply Normalization or Standardization of the dataset; the best practice would be to fit the trained models to normalized and standardized data and then compare their performances for the best results.

Normalization transforms features to a similar scale. The transformed value X_{new} is calculated as:

$$X_{new} = (X - X_{min})/(X_{max} - X_{min}) \quad 3.1$$

where, X is the actual value, X_{max} is the maximum value of the feature and X_{min} is the minimum value of the feature. This scales the range to [0, 1].

Standardization transforms the features by subtracting from mean and then dividing the result by standard deviation of the feature. The transformed value X_{new} can be presented as:

$$X_{new} = (X - X_{mean})/X_{std} \quad 3.2$$

Where, X is the actual value, X_{mean}n is the mean of all the values in the feature, and X_{std} is the standard deviation of the feature.

Our initial testing showed that standardization provided slightly better performance (lower RMSE). Therefore, standardization feature scaling was used before training the dataset for all the algorithms in this study.

3.5 Calibration Process

The calibration was framed as a supervised regression problem such that

$$X_{calibrated} = \Phi\{X_{LCS}, \underline{\mathbf{X}}\} \quad 3.3$$

Here $X_{calibrated}$ is the calibrated reading computed from the raw or pre-processed reading of the LCS (X_{LCS}) and $\underline{\mathbf{X}}$, that comprises of co-variate factors like T, RH, timestamped information and pollutant readings from the sensor array (e.g., other uncalibrated readings from the LCS array). Φ is the regression model whose parameters are derived from the training data to minimize the Mean Square Error (MSE) between the calibrated output and the ground truth received from the reference sensor.

For this study, we have considered two different Train Test Split (TTS) conditions. Here, the training set is a subset of the dataset. In TTS1, each of the datasets is split so that 90% of the data is used to train (and validate as discussed later) the calibration model while the remaining 10% data is used for evaluating the performance of the trained model. *This 90/10 split represents the scenario where a co-located LCS are being used as a backup in case the reference grade monitor is out of commission for a short period due to fault or maintenance*. In TTS2, the train/test split is 20/80. In TTS2, 20% of the data is used to train and validate the calibration model while the remaining 80% data is used for evaluating the performance of the trained model. *This emulates a scenario when a low-cost sensor is co-located with a reference sensor for a set period for calibration and afterwards deployed in the field for monitoring pollutants at locations where there is no AQM station is available*.

3.5.1 Training and Validation

As mentioned in the previous section, the regression model needs to be determined. Three types of calibration models have been prepared for this study, namely, MLR (or LR), Ensemble techniques and NN techniques. Although the MLR training and testing is fairly straightforward, a rigorous training, validation and testing method has been followed in this work for training and validating the regressors that construct the Ensemble and especially the NN based calibration models. The details of the calibration process are given here:

3.5.1.1 Calibration Steps for MLR Method

The following steps are conducted for the MLR method:

Step 1: The dataset is split into two parts: training and testing datasets.

Step 2: MLR (or LR) is performed on the training datasets by fitting a regressor model Φ between the LCS and reference sensors' training data.

Step 3: The regressor model Φ is used on the LCS testing dataset to get the calibrated (predicted) values for pollutant concentration.

Step 4: The final output is evaluated on the (unseen) testing dataset i.e., reference data, by computing the performance metrics.

3.5.1.2 Calibration Steps for Ensemble Method

The calibration models are prepared through training and validation process before testing them on the unseen testing dataset. All the regressors have hyperparameters that have been tuned on the relevant training datasets and tested on the corresponding testing sets. One way to make sure that the parameters are more generalized is through validation [196]. A k-fold (k =10) cross-validation has been implemented in this work. The following steps and Figure 3.5 provide a detailed description of this process:

Step 1: The dataset is split into two parts: training and testing datasets.

Step 2: The training is conducted using a ten-fold cross-validation, where the training dataset is divided into ten equal-sized parts. Each time nine out of the ten parts are used to perform grid search for hyperparameters tuning and then evaluated against the remainder 10th part (validation). The MSE (Mean Squared Error) is calculated on this holdout fold. This process is repeated ten times and the overall MSE is observed by averaging the ten MSE values from all the 10 folds. The best hyperparameter combination is found across all ten evaluations by selecting the one with the lowest overall MSE. Table 3.4 lists the hyperparameters that were tuned for the ensemble based algorithms.

Step 3: The best-performing model is now further trained using the entirety of the training dataset and a regressor model Φ is trained.

Step 4: The regressor model Φ is used on the LCS testing dataset to get the calibrated (predicted) values for pollutant concentration.

Step 5: The final output is evaluated on the (unseen) testing dataset by computing the performance metrics.

Table 3.4: List of hyperparameters that were tuned for ensemble based algorithms.

Algorithm	List of hyperparameters
RFR	maximum depth of the tree, maximum number of leaf nodes, number of trees in the forest
GBR	maximum depth of the individual regression estimators, minimum number of samples required to be at a leaf node, minimum number of samples required to split an internal node

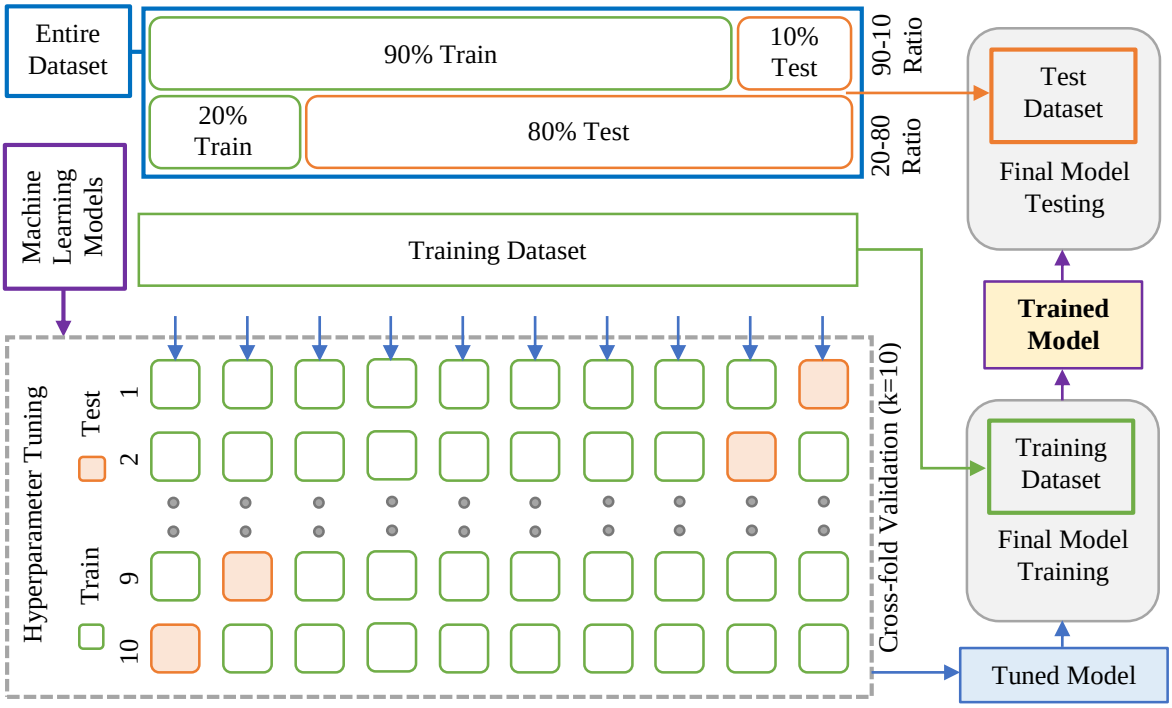


Figure 3.5: Details of k-fold cross-validation hyperparameter training.

3.5.1.3 Calibration Steps for NN Method

The NN methods have a grid search on top of the ten-fold cross-validation hyperparameter tuning. Each grid based on the combination of different model parameters has individual ten-fold cross-validation performed to tune the hyperparameters. The overall process is mostly similar to the ensemble methods in the first and the last three steps mentioned in section 3.5.1.2. However, the last step to select the final output for NN

based calibration also has an additional step compared to the ensemble ones. This calibration process is described below:

Step 1: The dataset is split into two parts: training and testing datasets.

Step 2: A grid search was performed on Neural Network based algorithms, namely MLP, LSTM and 1DCNN for the parameters presented in Table 3.5.

Table 3.5: Parameters for the grid search on Neural Network based algorithms.

Algorithm	Parameters for grid search
MLP	Number of hidden layers (1, 2, 3)
LSTM	Number of LSTM layers (1, 2, 3), number of time steps (1, 3, 5)
1DCNN	Number of 1D Convolution layers (1, 2, 3), lookback (1, 3, 5)

Step 3: For each grid, the training is conducted using a ten-fold cross-validation, where the training dataset is divided into ten equal-sized parts. Each time nine out of the ten parts are used to perform grid search for hyperparameters tuning and then evaluated against the remainder 10th part (validation). The MSE (Mean Squared Error) is calculated on this holdout fold. This process is repeated ten times and the overall MSE is observed by averaging the ten MSE values from all the 10 folds. The best hyperparameter combination is found across all ten evaluations by selecting the one with the lowest overall MSE. Table 3.6 lists the hyperparameters that were tuned for the neural network (NN) based algorithms. Similar process mentioned in Figure 3.5 at section 3.5.1.2 has been followed for calibration. However, instead of letting the training datasets to run for a set number of epochs, an early stopping method has been used during the training-validating stage. This method allowed the training to end once the model performance stopped improving on the validation set. MSE of the validation sets were monitored for each epoch and the training would stop when the MSE ceased to decrease by a certain tolerance threshold for a select number of epochs (patience). The model weights with the minimum MSE within that patience was taken as the final weight for that validation set.

Table 3.6: List of hyperparameters that were tuned for each NN based algorithm.

Algorithm	List of hyperparameters
MLP	number of neurons in the hidden layer, activation function in the hidden layer, dropout rate in dropout layer, learning rate of the optimizer, batch size
LSTM	number of units in the LSTM layers, activation function, dropout rate in dropout layers, learning rate of the optimizer, batch size
1DCNN	number of filters in the convolution layer, activation function in the convolution layer, kernel size, pool size in maxpooling layer, dropout rate in dropout layer, number of neurons in the dense layer, activation function in the dense layer, learning rate of the optimizer, batch size

Step 4: The best-performing model is now further trained using the entirety of the training dataset and a regressor model Φ is trained for each grid.

Step 5: The regressor model Φ is used on the LCS testing dataset to get the calibrated (predicted) values for pollutant concentration for each grid.

Step 6: RMSE values for each grid are compared between all grids (based on the number of parameters on the grid search). The grid parameter and model hyperparameter values with the lowest RMSE in the grid search based on top of a ten-fold cross-validation are selected as the final output.

3.6 List of Algorithms Applied

Six different algorithms have been used in this work, namely, LR/MLR (Multiple Linear Regression), RFR (Random Forest Regression), GBR (Gradient Boosting Regression), MLP (Multi-Layer Perceptron), LSTM (Long Short-Term Memory) and 1DCNN (One Dimensional Convolutional Neural Networks). Among these mentioned six algorithms, LR/MLR and MLP are the two most commonly used ones found from the extensive literature review (Chapter 2). Also, RFR, which is a type of ensemble machine learning technique, has been used on many occasions and was reported to be a good option for LCS calibration. Different types of RNN has been also used on some studies of LCS calibration. LSTM is a type of RNN and has been used in this study. The other two used algorithms in this work are GBR and 1DCNN. GBR is a type of ensemble technique similar to RFR and 1DCNN (like MLP and LSTM) falls under the category of

Neural Networks. Our investigations show that GBR and 1DCNN are have been rarely used in the previous studies and have not been properly explored for LCS. In addition, these previous studies lack performance benchmarking against other frequently used algorithms, making it impossible to compare their efficiency. Chapter 2 provides a detailed literature review of relevant algorithms summarized in Table 2.7.

Thus, these six applied algorithms cover the broader category of different machine learning techniques mentioned in the literature, where some of them being frequently used while others need more investigations. All these algorithms are briefly described in the following sections.

3.6.1 LR/MLR

Linear regression (LR) or multiple linear regression (MLR) is the single most commonly used technique to calibrate LCS (see Section 2.4.2 and Table 2.7). The regression method is used for finding the linear relationship between target and one or more variables [30, 31]. This calibration technique is popular due to its simplicity. Two types of linear regression are found in the literature for calibrating air pollutant LCS – simple (or ordinary) [11, 161] and multiple linear regression [18, 49, 65, 148, 163].

Linear regression provides a linear model Φ, e.g., a model that assumes a linear relationship between the input variables (x_1, x_2…, x_n) and the single output variable (y). More specifically, that y can be calculated from a linear combination of the input variables (x_1, x_2…, x_n). When there is only one input variable (x) e.g., raw response from a LCS, the method is referred to as simple linear regression. On the other hand, when there are multiple input variables e.g. temperature, humidity, other gas pollutant values, the literature often refers to the method as multiple linear regression (MLR) [65]. Of the different techniques that can be used to train a linear regression equation from data, the most common of which is called Ordinary Least Squares [11, 30, 31]. It is common to therefore refer to a model prepared this way as Ordinary Least Squares Linear Regression or just Least Squares Regression. The expression of the regression model Φ in a multiple linear regression calibration model would be:

$$y = B_0 + B_1 x_1 + B_2 x_2 + ... + B_n x_n \quad (3.4)$$

where, B_0 is the slope and, B_1, B_2, … , B_n represent specific values used for the coefficients of each input parameter (x_1, x_2…, x_n). The input parameters for LCS

calibration can represent raw LCS data (working and/or auxiliary electrode data), temperature and relative humidity etc. These coefficients are calculated by minimizing the Mean Squared Error (MSE) between the values of the variable being observed in the input dataset observed (dependent variables x_1, x_2…, x_n) and the output of the (linear) function of the independent variable y.

3.6.2 RFR

Random Forest Regression (RFR) algorithm is a supervised learning ensemble learning method for regression [197]. Ensemble learning method is powerful and accurate as it combines predictions using multiple machine learning algorithms to predict more accurately than by using a single learning model. As seen in Chapter 2, RFR is a popular calibration model for LCS and RFR based calibration techniques have performed well for LCSs. For example, Borrego *et al.* implemented calibration technique on LCS that measures gas and particulate matter (PM) sensors [26, 34]. Also, Cordero *et al.* [65] used RFR to improve the accuracy of NO_2 sensors for field deployment. Similarly, Bigi *et al.* [67], Bagkis *et al.* [69] and Malings *et al.* [71] also implemented RFR on field deployed electrochemical gas sensors and found satisfactory performance. All these performance improvements were claimed to be better in terms of lower RMSE values than just using a simple MLR as calibration algorithm. However, the accuracy improvements were not benchmarked with respect to other commonly used methods. Zimmerman *et al.* [68] benchmarked RFR and MLR on multiple gas sensors for both lab and field deployments and found that the RFR performance was better than MLR. One of the main reasons for RFR being a suitable candidate in all these examples is its ability to account for the pollutant cross-sensitivity [68]. Therefore, we decided to utilize RFR as one of the ML-based calibration techniques.

Although RFR performs well with non-linear relationships, overfitting (machine learning model giving accurate prediction for training data but not for new data [198]) may easily occur [199, 200]. To overcome this, a rigorous training and validation process is followed for these algorithms as mentioned in Section 3.5.1.2. Figure 3.6 shows the typical example of an Ensemble learning model for RFR.

We have used the sklearn module for training our RFR model by using RandomForestRegressor function. Following the steps mentioned earlier for calibrating

ensemble models (Section 3.5.1.2), we have tuned the following parameters during the training-validation process using MSE as loss function:

3.6.2.1 n_estimator

This represents the number of decision trees running in the regression model. A large number of trees can sometime reduce overfitting in RFR as increasing the number of trees generally reduces model error at the cost of a longer training time and higher computational power.

3.6.2.2 max_depth

This parameter sets the maximum depth of each tree. More information about the data is taken into account with larger depth of the tree. Thus possible number of feature or value combinations can be increased by increasing the tree depth. In an individual tree this increase can cause overfitting, however in RFR, it is harder to overfit due to the way the ensemble is built. But it is still possible to overfit for a very large depth value. So this parameter should be set to a reasonable value depending on the number of features of your tree.

3.6.2.3 max_leaf_nodes

This parameter controls the growth of the trees by limiting the maximum number of leaf nodes. If not defined, then unlimited number of leaf nodes can be produced in the training.

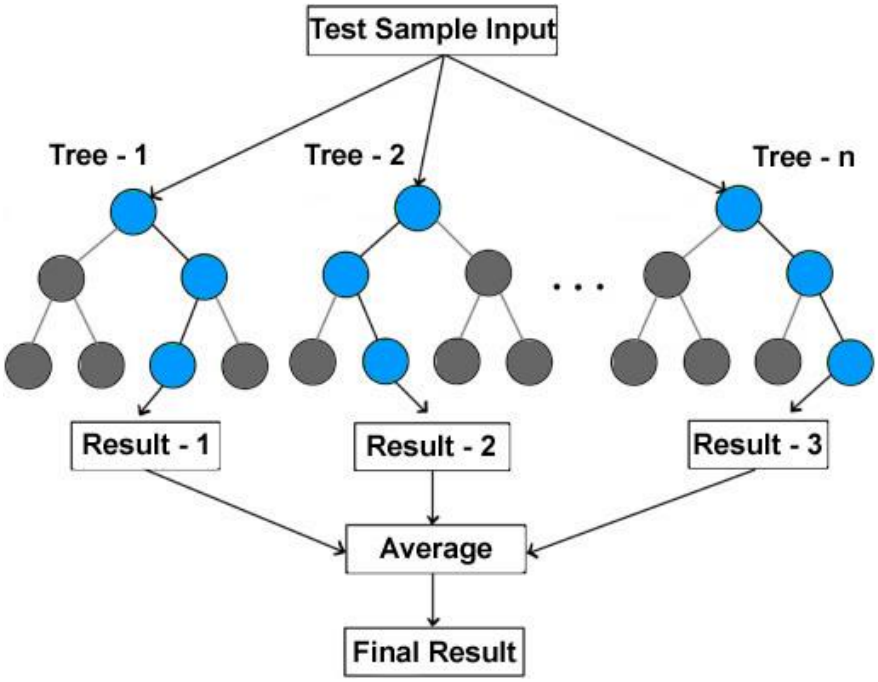


Figure 3.6: General structure of ensemble type algorithms (RFR and GBR).

3.6.3 GBR

Gradient Boosting Regression (GBR) is based on boosting algorithms and is considered to be more efficient as the algorithm deals with bias variance trade-off by controlling both aspects (bias & variance) [201]. GBR is a sequential technique which uses the principles of ensemble models. Improved prediction accuracy has been obtained using boosting based regression [202]. Figure 3.6 shows the typical example of an Ensemble learning model for GBR.

GBR has also not been widely utilized for gas sensor calibration and the performance of GBR was not benchmarked against sophisticated techniques like NNs (see Section 2.4.2 and Table 2.7). Training and validation process mentioned in Section 3.5.1.2 has been used to calibrate the LCS data and the following parameters have been tuned in this study for GBR during the training and validation stage:

3.6.3.1 max_depth

This represents the maximum depth of a tree. A higher depth can allow model to learn relations very specific to a particular sample and thus control over-fitting.

3.6.3.2 min_sample_split

It is the minimum number of observations (or samples) that are required in a node to be considered for splitting. This parameter is for controlling the model over-fitting. A higher value can prevent the model from learning relations which might be highly specific to a particular sample selected for a tree. However, too high values can lead to under-fitting of the model.

3.6.3.3 min_sample_leaf

This represents the minimum observations (or samples) required in a terminal node or leaf. It is used to control over-fitting similar to min_samples_split.

3.6.4 MLP

Artificial neural networks (ANN) are often just called neural networks (NN) or multi-layer perceptron (MLP). MLP is the most commonly used NN-based calibration technique found in the literature (see Section 2.4.2 and Table 2.7). The learning capability of MLPs comes from the multi-layered structure of the networks [203]. The building block of a neural network is the artificial neuron. It is a simple computational unit which has weighted input signal and produce an output signal by using an activation function [204]. This activation function governs the threshold at which the neuron is activated and strength of the output signal. In terms of LCS, the neurons in the MLP learn the representation of different input variables in training data and relate it to the output variable i.e., the ground truth reference data [17, 48, 50, 57, 58, 174]. Figure 3.7 represents a simple illustration of one of the MLP structures used in this study. Here, the LCS data such as raw gas concentrations, temperature, relative humidity is used as the input parameters and the reference gas concentration from the ground truth sensor is used as the target output to train the weights in the hidden layers. Different MLP structures have been studied in this work by changing the number of neurons in the hidden layer, activation function in the hidden layer, dropout rate in dropout layer, learning rate of the optimizer, batch size etc. Table 3.7 represents the grid search table used for MLP algorithms. Each grid for different number of hidden layers was tuned for hyperparameters according to the calibration process for this algorithm is mentioned in

Section 3.5.1.3. The hidden layer and hyperparameters with the lowest RMSE was selected as the final results for MLP algorithms.

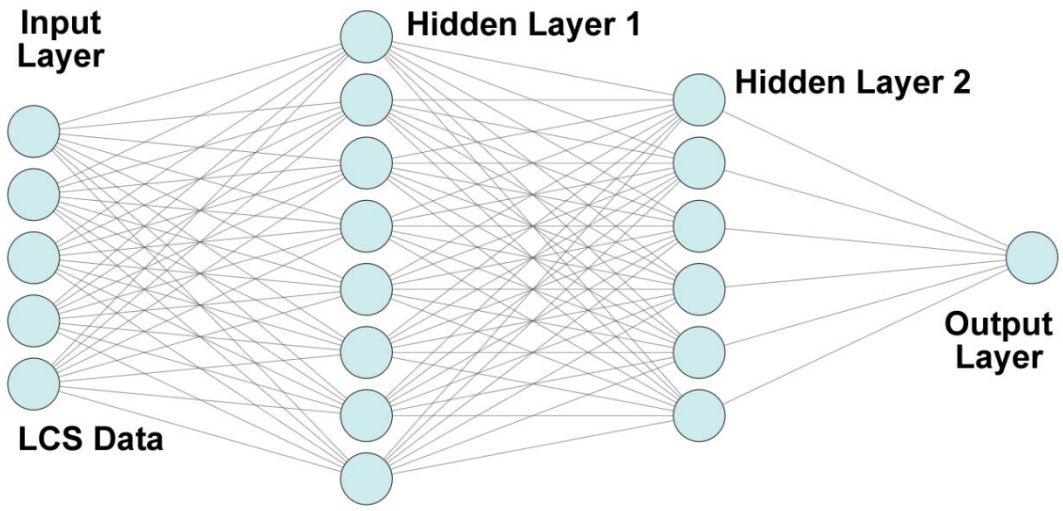


Figure 3.7: Example of a MLP architecture. Please note that the structure will vary depending on the many parameters are determined through grid search and tuning.

Table 3.7: Grid search used in addition to the k-fold cross-validation for MLP.

Number of hidden layers	**Hyperparameters selected after the k-fold cross-validation**	**RMSE from the unseen testing dataset**
1	Hyperparameters A	RMSE value A
2	Hyperparameters B	RMSE value B
3	Hyperparameters C	RMSE value C

3.6.5 LSTM

Many researchers have used Recurrent Neural Networks to calibrate LCS data (see Table 2.7). Unlike other NNs that mostly use current data, Recurrent Neural Networks (RNN) model the historical time series behaviour present in the dataset. Recurrent Neural Networks (RNN) have been used by Sheik *et al.* [169] Wang *et al.* [178] and Fonollosa *et al.* [76] for calibrating MOX-based LCSs under laboratory conditions. Esposito *et al.* [57, 58] studied multiple calibration techniques including RNN on different EC-based LCSs and compared their performance. It should be noted that RNN

models face two issues during calibration: firstly, time lag must be determined in advance which requires a considerable number of experiments to identify them, and secondly, these RNNs fail to capture long time dependencies in the training dataset. Thus Long Short-Term Memory networks (LSTMs) are introduced [77]. LSTMs are basically an extension for RNN which extends the memory to learn from important changes that have large time gaps in between [77, 205]. Thus, LSTM is able to process longer time series information than RNN in terms of data processing, error correction and forecasting [75, 206-208]. As a result, it has been used for forecasting and predicting air quality gas sensor data by different research groups [75, 86, 206, 207, 209]. Apart from gas pollutant prediction and calibration, RNN as Long Short-Term Memory (LSTM) was also used by Li *et al.* [206] and Athira *et al.* [75] to improve the accuracy of the field deployed PM sensors. All these studies have found LSTM to generate better in predicting sensor outputs given that the training datasets are large enough to capture the time dependent relationships among the input parameters if they are there. Section 3.5.1.3 have the details of calibration method for LSTM in this study. Table 3.8 represents the grid search table used for LSTM algorithm. This grid is based on the number of LSTM layers used in the model and the number of lag observations as time steps between 1, 3 and 5.

Each grid for different number of LSTM layers and lag observations was tuned for hyperparameters according to the calibration process for this algorithm is mentioned in Section 3.5.1.3. The grid with the lowest RMSE was selected as the final results for LSTM algorithms.

Table 3.8: Grid search table used in addition to the k-fold cross-validation for LSTM.

Number of lag observations	**Number of LSTM layers**		
	1	2	3
1	RMSE value	RMSE value	RMSE value
3	RMSE value	RMSE value	RMSE value
5	RMSE value	RMSE value	RMSE value

3.6.6 1DCNN

Convolutional Neural Networks (CNNs) have become the most popular machine learning technique during the last decade [210]. Deep CNNs can be generally defined as feed forward Artificial Neural Networks (ANNs) with many hidden layers that provides them the ability to learn complex relationships from the training data [211]. The conventional CNNs (a.k.a. 2DCNNs) are mainly designed to process two dimensional (2D) data e.g., videos and images [211]. This structure can be modified and developed as 1DCNN to deal with one dimensional signals [212-214]. The 1DCNN algorithms have less computational complexity, compact structure (1-2 hidden CNN layers), less time consuming to train and thus suitable for real-time low-cost applications to their 2D counterparts [211]. 1DCNN are mainly constructed using two parts: in the first part, 1D convolutions, activation function and pooling occur, and in the second part, fully connected dense layer e.g., MLP are inserted [211]. Literature review from Chapter 2 shows that among the NN-based techniques, 1DCNN has not been well investigated for low-cost gas sensor calibration.

Table 3.9 represents the grid search table used for 1DCNN algorithm. This grid is based on the number of 1DCNN layers used in the model and the lookback observations as considering previous observations steps between 1, 3 and 5. Each grid for different number of 1DCNN layers and lookback observations was tuned for hyperparameters according to the calibration process for this algorithm is mentioned in Section 3.5.1.3. The lowest RMSE value was selected as the final results for 1DCNN algorithms from this grid search performed over the k-fold cross-validation. Figure 3.8 shows an example of the 1DCNN structure used in this study.

Table 3.9: Grid search used in addition to the k-fold cross-validation for 1DCNN.

Feedback observations	**Number of 1DCNN layers**		
	1	2	3
1	RMSE value	RMSE value	RMSE value
3	RMSE value	RMSE value	RMSE value
5	RMSE value	RMSE value	RMSE value

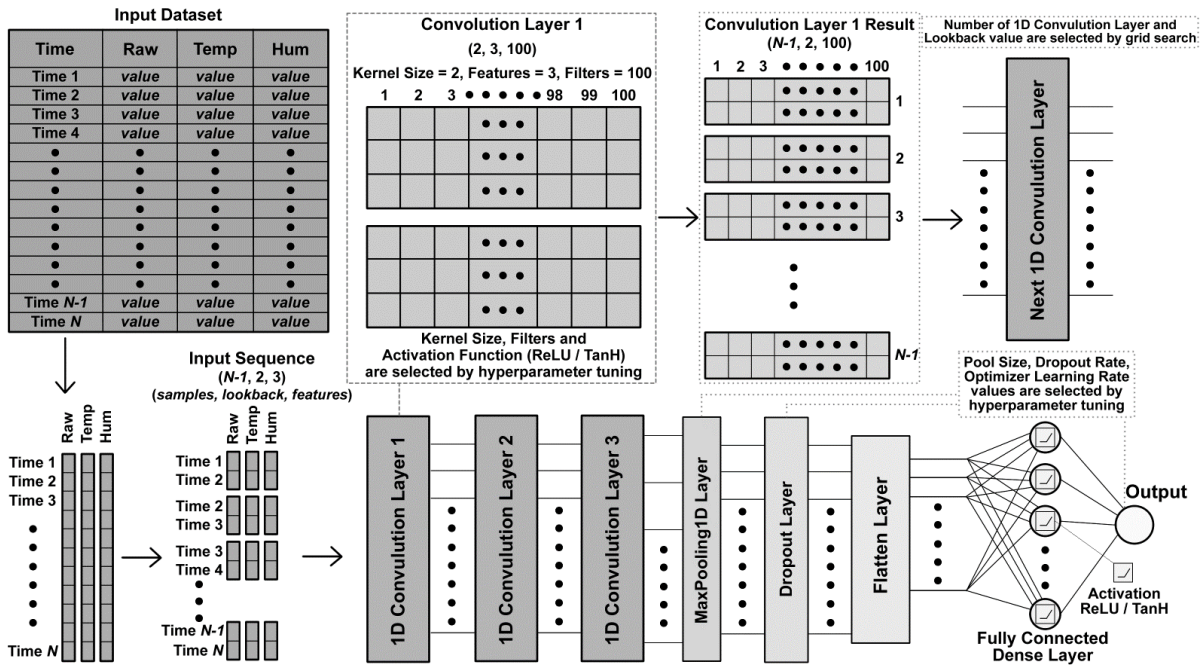


Figure 3.8: An example of a 1DCNN model developed for this study. Please note that many parameters are determined through grid search and tuning. Therefore, they vary depending on the data set, scenario (input variables used), and time split.

3.6.7 Parameters Tuned for NN-based Algorithms

The following parameters have been tuned for the three Neural Network based deep learning algorithms (MLP, LSTM and 1DCNN):

3.6.7.1 Number of Hidden Layers

The hidden layer is the layer between the input and output layers. This fundamental structure makes deep learning networks often referred as a “black box” as their predictions in each step of the network are not traceable normally. There is no final number on how many nodes (hidden neurons) or hidden layers one should use in a NN. So this number may vary from one type of problem to other. Generally, one hidden layer will work with most simple problems while two layers are suitable for more non-linear ones. Large number of nodes within a layer can increase accuracy and fewer number of nodes may cause underfitting, however, too many nodes may cause over-fitting of the model.

3.6.7.2 Number of Nodes

In a hidden layer each neuron receives input from all neurons in the previous layer. Number of nodes or units in a hidden layer dictates the computational power of a network. Large number of nodes within a layer can increase accuracy and fewer number of nodes may cause under-fitting, however, too many nodes may cause over-fitting of the model. Increasing nodes in these layers or adding a greater number of hidden layers may result in poor validation accuracy. However, this can be resolved by the next parameter i.e., dropout.

3.6.7.3 Dropout Rate

Dropout layer helps to avoid overfitting in training by bypassing randomly selected neurons. Thus, the sensitivity to specific weights of the individual neurons can be reduced to a certain level. It is usually used with the input layer or hidden layers. The dropout layer shouldn't be used with the output layer as this may mess up the output from the model. Sometime overfitting in a complex structure can be addressed by adding dropout layers. The rate of the dropout should be tuned in the training-validation stage.

3.6.7.4 Activation Function

Activation functions are used to introduce non-linearity to models. The choice of activation function depends on the application. For regression type problems with non-linear relationship present in the samples, *relu* or *tanh* activation functions are most popular and have been used in this study.

3.6.7.5 Learning Rate

Learning rate defines how quickly the network changes its parameters. It has a small positive value, mostly between 0.0 and 0.1. A higher learning rate accelerates the learning, but the model may not converge (it is the state during training stage where the loss value settles to within an error range around the expected value), or even diverge. Conversely, a lower rate can slow down the learning drastically but will allow the model to converge smoothly.

3.7 Performance Metrics

Several performance metrics have been used in this study to benchmark and evaluate the different calibration models. These metrics in various ways measure the residuals or errors i.e., deviations of calibrated output of the LCS ($X_{calibrated}$) from the ground truth ($X_{reference}$) for the test data (10% or 80% of every dataset, depending on the split) that has never been used for training.

Root Mean Squared Error (RMSE) which is the standard deviation of the residuals and commonly used as a performance metric for sensor calibration [14, 215-218] were used where

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}\left[X_{calibrated} - X_{reference}\right]^2} \quad 3.6$$

Here *N* is the number of samples in the relevant test data set.

Another metrics utilized is the Coefficient of Determination (R^2) which is the goodness of fit in a regression analysis [7, 209, 219]. It is computed as

$$R^2 = 1 - \frac{\sum_{i=0}^{N-1}\left[X_{calibrated} - X_{reference}\right]^2}{\sum_{i=0}^{N-1}\left[X_{reference} - mean(X_{reference})\right]^2} \quad 3.7$$

While the appropriateness of R^2 to determine the efficacy of non-linear regressors has been questioned, it is still commonly used within the discipline of air pollutant measurement (e.g., see [30, 31, 68, 218, 219]).

In some instances, we have also plotted Cumulative Distribution Function (CDF) [220] of absolute errors, $abs\left[X_{calibrated} - X_{reference}\right]$, for more detailed investigation.

Target diagrams [31, 68] were constructed for visualizing the calibration models (please see Figure 3.9). The y axis represents the Mean Bias Error (MBE) normalized by the standard deviation of the ground truth so that

$$MBE = mean(X_{calibrated}) - mean\left(X_{reference}\right) \quad 3.8$$

$$Normalised\ MBE = \frac{MBE}{\sigma_{reference}} \quad 3.9$$

Here, $\sigma_{reference}$ is the standard deviation of the ground truth for the relevant test data set. The x axis of the diagram represents the normalized unbiased estimate of the RMSE, the Centered RMSE (CRMSE) where

$$CRMSE = \sqrt{RMSE^2 - MBE^2} \tag{3.10}$$

$$Normalised\ CRMSE = \frac{CRMSE}{\sigma_{reference}} \tag{3.11}$$

Please note that the normalized CRMSE is multiplied by $sign\{\sigma_{calibrated} - \sigma_{reference}\}$ to produce the target diagrams with $\sigma_{calibrated}$ being the standard deviation of the calibrated data for the relevant test data set. The general description of a target diagram (see Figure 3.9) is given below:

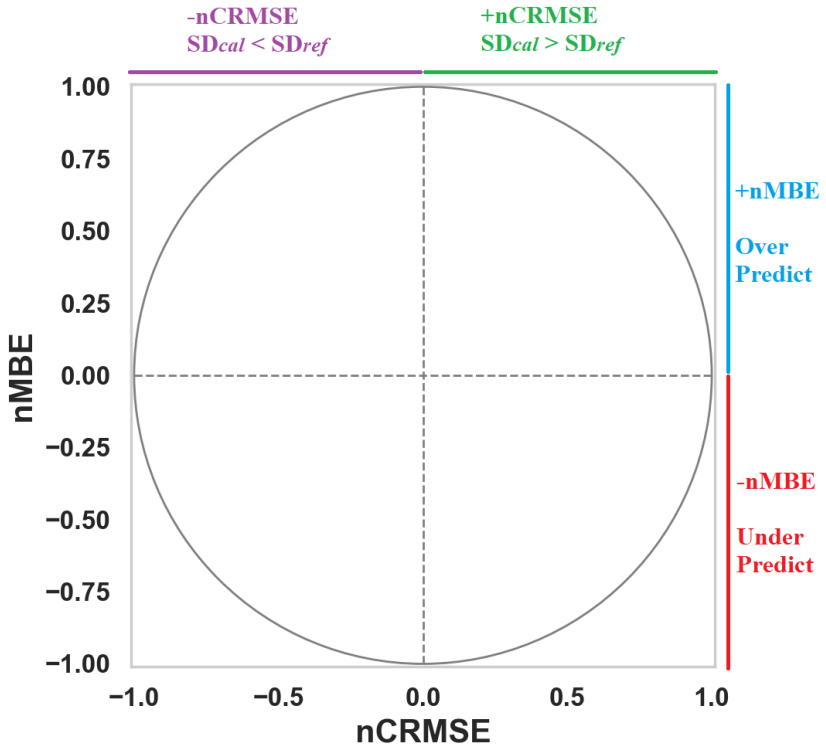


Figure 3.9: General description of a target diagram.

1. When points lie within the unit circle (radius = 1), it indicates that the variance of the residuals is smaller than the variance of the reference measurements. It is an essential characteristic of a functional calibration model [31], indicating that the variability of the dependent variable (calibrated output) is explained by the independent variable (the reference data) and not the residual [53].

2. The distance from the origin to a particular point measures the normalized RMSE (RMSE/$\sigma_{reference}$) for that regressor. Therefore, points closer to the origin a point is more accurate the corresponding model is.
3. Left and right planes of the y-axis show the difference of Standard Deviation (SD) between calibrated and reference samples. Points on the right side of the y-axis mean that the SD of the calibrated data is larger than the SD of the reference data and vice versa [31, 53].
4. Above and below x-axis show over or underestimation of the calibration; where, points above the x-axis will be overestimating and points below the x-axis will be underestimating [31, 53].

CHAPTER 4

RESULTS & DISCUSSIONS: ALGORITHM BENCHMARKING

This chapter presents a comprehensive benchmarking of several machine-learning algorithms with respect to their performance for calibrating LCS. We demonstrate that algorithms like 1DCNN and GBR which have not been extensively used for low-cost gas sensor calibration provide consistently accurate results and thus should be considered as viable alternatives.

4.1 CO and NO_2 Results

The algorithms are used for both CO and NO_2 gas calibration for the three Datasets introduced in Chapter 3. Both CO and NO_2 are important components of the Air Quality Index (AQI) [156], and both the raw (unchanged electrode data from the LCS) and reference data (from high-end accurate sensors) are available for all three deployments. In the previous chapter, we have already established the importance of temperature and humidity correction as well as the need to mitigate cross-sensitivity from the LCS raw responses. These deployments include additional pollutant data that enable us to address the cross-sensitivity issues of LCS. Furthermore, both temperature and relative humidity data are available for all three setups, allowing us to mitigate the effects of temperature and relative humidity. However, the multisensory LCS devices and reference sensors have missing samples and readings of select time instants from each dataset. If any pollutant data from the LCSs or the ground truth data were missing, they have been removed before training the models.

4.2 Calibration Scenarios

Based on the literature, it can be summarized that the calibration of LCS has been done roughly in three ways regardless of the algorithm used. One, by only using the raw LCS data to calibrate with respect to the reference sensor [11, 30, 31, 157]. This is essentially performing an offset and gain calibration along with trying to model some non-linearity in the LCS response. Secondly, by introducing temperature and relative

humidity values with the raw data during calibration to correct their effects on the sensor response [18, 49, 68, 103, 120]. Lastly, including other gas concentrations as input parameters on top of the existing scenario two to account the cross-sensitivity issue [17, 36, 48, 50, 58, 94]. In this work, all three scenarios were thoroughly investigated for all three datasets for both gases, CO and NO_2. We have named these scenarios (one, two and three) as SC1, SC2 and SC3, respectively. The same calibration method has been used for all six algorithms in these three scenarios, the details of which have been discussed in the previous chapter. All three datasets have been processed through data cleaning, feature engineering, feature scaling, and the training and validation process as outlined in Chapter 3. Different input features have been selected for different scenarios for the three datasets illustrated in Table 4.1 – 4.6.

Table 4.1: Input feature selection for scenarios in Dataset 1 for CO.

Scenario	**Selected Input Features**
SC1	CO (raw)
SC2	CO (raw), T, RH
SC3	CO (raw), NO_2 (raw), NO_X (raw), O_3 (raw), NMHC (raw), T, RH

Table 4.2: Input feature selection for scenarios in Dataset 1 for NO_2.

Scenario	**Selected Input Features**
SC1	NO_2 (raw), NO_X (raw)
SC2	NO_2 (raw), NO_X (raw), T, RH
SC3	NO_2 (raw), NO_X (raw), CO (raw), O_3 (raw), NMHC (raw)

Table 4.3: Input feature selection for scenarios in Dataset 2 for CO.

Scenario	**Selected Input Features**
SC1	CO_we (raw), CO_ae (raw)
SC2	CO_we (raw), CO_ae (raw), T, RH
SC3	CO_we (raw), NO_2_we (raw), O_3NO_2_we (raw), CO_ae (raw), NO_2_ae (raw), O_3NO_2_ae (raw), T, RH

Table 4.4: Input feature selection for scenarios in Dataset 2 for NO_2.

Scenario	Selected Input Features
SC1	NO_2_we (raw), NO_2_ae (raw)
SC2	NO_2_we (raw), NO_2_ae (raw), T, RH
SC3	NO_2_we (raw), CO _we (raw), O_3NO_2_we (raw), NO_2_ae (raw), CO _ae (raw), O_3NO_2_ae (raw), T, RH

Table 4.5: Input feature selection for scenarios in Dataset 3 for CO.

Scenario	Selected Input Features
SC1	CO_we (raw), CO_ae (raw)
SC2	CO_we (raw), CO_ae (raw), T, RH
SC3	CO_we (raw), NO_2_we (raw), O_x_we (raw), CO_ae (raw), NO_2_ae (raw), O_x_ae (raw), T, RH

Table 4.6: Input feature selection for scenarios in Dataset 3 for NO_2.

Scenario	Selected Input Features
SC1	NO_2_we (raw), NO_2_ae (raw)
SC2	NO_2_we (raw), NO_2_ae (raw), T, RH
SC3	NO_2_we (raw), CO_we (raw), O_x_we (raw), NO_2_ae (raw), CO _ae (raw), O_x_ae (raw)

EC sensors are based on electrochemical cells that generate current in proportion to target gas concentration [221]. The driver circuit that facilities the measurement of the target gas consists of two major electrodes, namely working electrode (we) and auxiliary electrode (ae) [222]. Depending on the hardware of the EC sensor, the target gas concentration information may come from one or both electrodes [14]. In this study, both electrode datasets were available for datasets 1 and 2 and were utilized for calibration.

4.3 Performance Comparison

Tables 4.7 – 4.12 show the performance of various algorithms for CO and NO_2 for the three datasets. Please note the top three performances have been marked based on RMSE values: First Second Third. Figures 4.1 – 4.12 illustrate the comparison of RMSE values between different scenarios for all algorithms.

Table 4.7: The accuracy of CO calibration algorithms in terms of R^2 and RMSE (in ppm) for Dataset 1.

TTS	Scenario	Performance Metric	Algorithm					
			LR/MLR	RFR	GBR	MLP	LSTM	1DCNN
TTS1	SC1	R^2	0.803	0.806	0.805	0.810	0.808	0.811
		RMSE	0.554	0.554	0.556	0.551	0.545	0.541
	SC2	R^2	0.812	0.837	0.838	0.833	0.838	0.841
		RMSE	0.543	0.508	0.506	0.526	0.501	0.502
	SC3	R^2	0.905	0.924	0.923	0.923	0.925	0.924
		RMSE	0.384	0.346	0.349	0.350	0.344	0.344
TTS2	SC1	R^2	0.767	0.770	0.771	0.768	0.780	0.781
		RMSE	0.613	0.609	0.609	0.613	0.600	0.598
	SC2	R^2	0.779	0.781	0.781	0.786	0.798	0.805
		RMSE	0.597	0.594	0.594	0.588	0.572	0.564
	SC3	R^2	0.882	0.899	0.897	0.896	0.900	0.904
		RMSE	0.437	0.404	0.409	0.415	0.405	0.396

The top three performances have been marked based on RMSE values: First Second Third

Table 4.8: The accuracy of NO_2 calibration algorithms in terms of R^2 and RMSE (in ppb) for Dataset 1.

TTS	Scenario	Performance Metric	Algorithm					
			LR/MLR	RFR	GBR	MLP	LSTM	1DCNN
TTS1	SC1	R^2	0.484	0.633	0.628	0.620	0.631	0.633
		RMSE	17.786	14.996	15.086	15.476	15.187	15.025
	SC2	R^2	0.710	0.820	0.813	0.785	0.825	0.835
		RMSE	13.327	10.502	10.709	11.587	10.366	10.059
	SC3	R^2	0.785	0.871	0.865	0.855	0.869	0.886

		RMSE	11.486	8.886	9.069	9.459	8.968	8.445
TTS2	SC1	R^2	0.489	0.601	0.597	0.617	0.607	0.623
		RMSE	17.206	15.201	15.281	14.933	15.155	14.874
	SC2	R^2	0.677	0.763	0.760	0.758	0.779	0.790
		RMSE	13.681	11.711	11.776	12.189	11.371	11.123
	SC3	R^2	0.744	0.817	0.806	0.825	0.811	0.831
		RMSE	12.187	10.291	10.593	10.082	10.516	9.894

The top three performances have been marked based on RMSE values: First Second Third

Table 4.9: The accuracy of CO calibration algorithms in terms of R^2 and RMSE (in ppm) for Dataset 2.

TTS	**Scenario**	**Performance Metric**	**Algorithm**					
			LR/MLR	**RFR**	**GBR**	**MLP**	**LSTM**	**1DCNN**
TTS1	SC1	R^2	0.507	0.837	0.842	0.713	0.837	0.841
		RMSE	0.305	0.175	0.172	0.252	0.175	0.173
	SC2	R^2	0.658	0.907	0.914	0.755	0.909	0.918
		RMSE	0.254	0.132	0.127	0.219	0.130	0.124
	SC3	R^2	0.710	0.910	0.923	0.868	0.924	0.927
		RMSE	0.234	0.129	0.120	0.204	0.119	0.117
TTS2	SC1	R^2	0.443	0.784	0.795	0.675	0.792	0.789
		RMSE	0.300	0.187	0.182	0.238	0.183	0.185
	SC2	R^2	0.616	0.865	0.869	0.713	0.870	0.871
		RMSE	0.249	0.147	0.145	0.217	0.147	0.145
	SC3	R^2	0.674	0.873	0.877	0.842	0.890	0.887
		RMSE	0.229	0.143	0.141	0.188	0.134	0.136

The top three performances have been marked based on RMSE values: First Second Third

Table 4.10: The accuracy of NO_2 calibration algorithms in terms of R^2 and RMSE (in ppb) for Dataset 2.

TTS	Scenario	Performance Metric	Algorithm					
			LR/MLR	RFR	GBR	MLP	LSTM	1DCNN
TTS1	SC1	R^2	0.495	0.518	0.526	0.514	0.525	0.527
		RMSE	8.612	8.415	8.343	8.447	8.357	8.351
	SC2	R^2	0.559	0.658	0.678	0.617	0.662	0.661
		RMSE	8.044	7.089	6.879	7.594	7.080	7.045
	SC3	R^2	0.649	0.741	0.798	0.733	0.763	0.762
		RMSE	7.183	6.193	5.453	6.432	5.896	6.030
TTS2	SC1	R^2	0.499	0.506	0.507	0.521	0.522	0.518
		RMSE	8.480	8.434	8.427	8.462	8.301	8.360
	SC2	R^2	0.550	0.618	0.618	0.612	0.624	0.625
		RMSE	8.037	7.419	7.412	7.523	7.366	7.338
	SC3	R^2	0.633	0.723	0.735	0.710	0.724	0.730
		RMSE	7.256	6.303	6.184	6.489	6.340	6.242

The top three performances have been marked based on RMSE values: First Second Third

Table 4.11: The accuracy of CO calibration algorithms in terms of R^2 and RMSE (in ppm) for Dataset 3.

TTS	Scenario	Performance Metric	Algorithm					
			LR/MLR	RFR	GBR	MLP	LSTM	1DCNN
TTS1	SC1	R^2	0.920	0.919	0.926	0.926	0.922	0.926
		RMSE	0.075	0.075	0.072	0.073	0.074	0.072
	SC2	R^2	0.948	0.970	0.972	0.965	0.971	0.973
		RMSE	0.060	0.046	0.044	0.049	0.045	0.044
	SC3	R^2	0.958	0.974	0.980	0.971	0.978	0.979
		RMSE	0.054	0.043	0.038	0.049	0.039	0.038
TTS2	SC1	R^2	0.895	0.890	0.890	0.897	0.894	0.901
		RMSE	0.080	0.082	0.082	0.082	0.080	0.079
	SC2	R^2	0.927	0.933	0.936	0.936	0.937	0.937

		RMSE	0.067	0.064	0.063	0.063	0.062	0.062
	SC3	R^2	0.941	0.941	0.954	0.957	0.954	0.967
		RMSE	0.060	0.060	0.053	0.056	0.053	0.049

The top three performances have been marked based on RMSE values: First Second Third

Table 4.12. The accuracy of NO_2 calibration algorithms in terms of R^2 and RMSE (in ppb) for Dataset 3.

TTS	**Scenario**	**Performance Metric**	**Algorithm**					
			LR/MLR	**RFR**	**GBR**	**MLP**	**LSTM**	**1DCNN**
TTS1	SC1	R^2	0.734	0.762	0.761	0.750	0.765	0.761
		RMSE	7.015	6.634	6.644	6.898	6.635	6.768
	SC2	R^2	0.792	0.873	0.877	0.853	0.886	0.891
		RMSE	6.206	4.849	4.776	5.517	4.711	4.567
	SC3	R^2	0.806	0.889	0.910	0.873	0.943	0.947
		RMSE	5.995	4.549	4.079	4.916	3.317	3.219
TTS2	SC1	R^2	0.779	0.787	0.778	0.756	0.797	0.803
		RMSE	6.433	6.323	6.453	7.493	6.181	6.128
	SC2	R^2	0.822	0.867	0.857	0.857	0.874	0.888
		RMSE	5.766	4.982	5.168	5.399	5.073	4.585
	SC3	R^2	0.828	0.881	0.890	0.852	0.914	0.919
		RMSE	5.685	4.737	4.541	5.582	4.134	3.930

The top three performances have been marked based on RMSE values: First Second Third

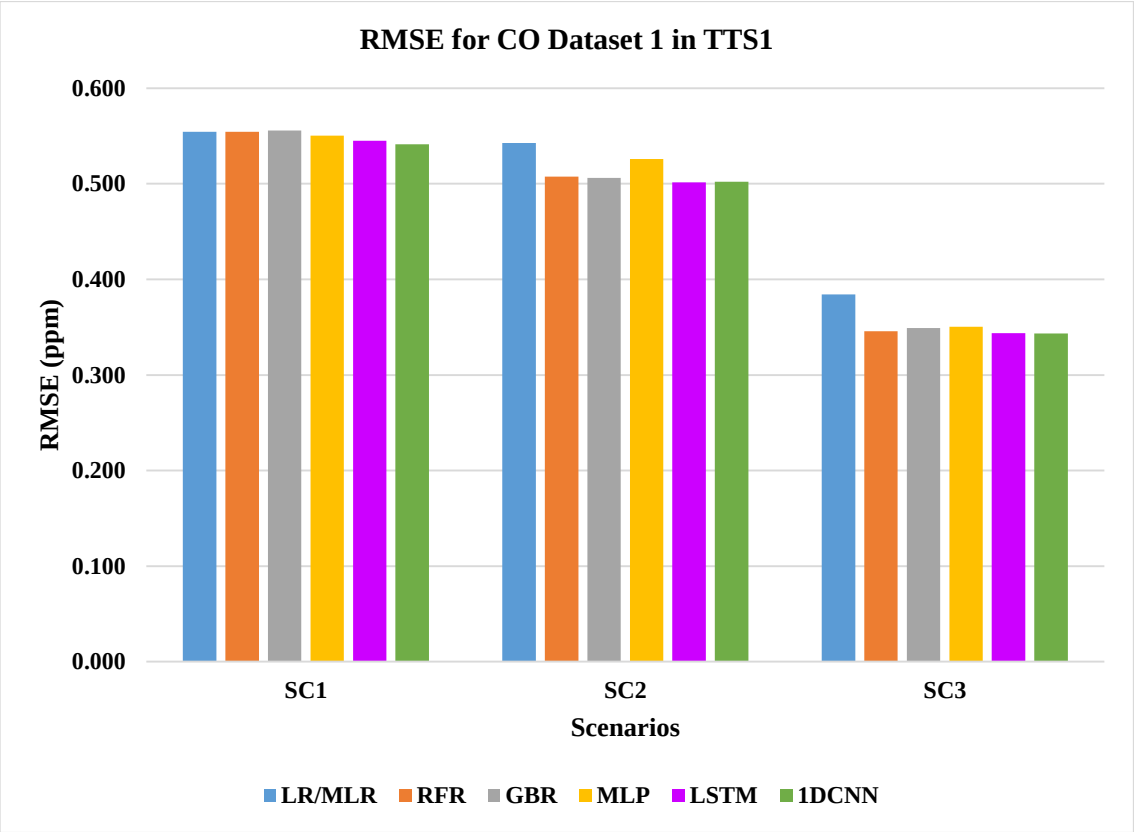


Figure 4.1: RMSE comparison between different scenarios and algorithms for CO Dataset 1 in TTS1.

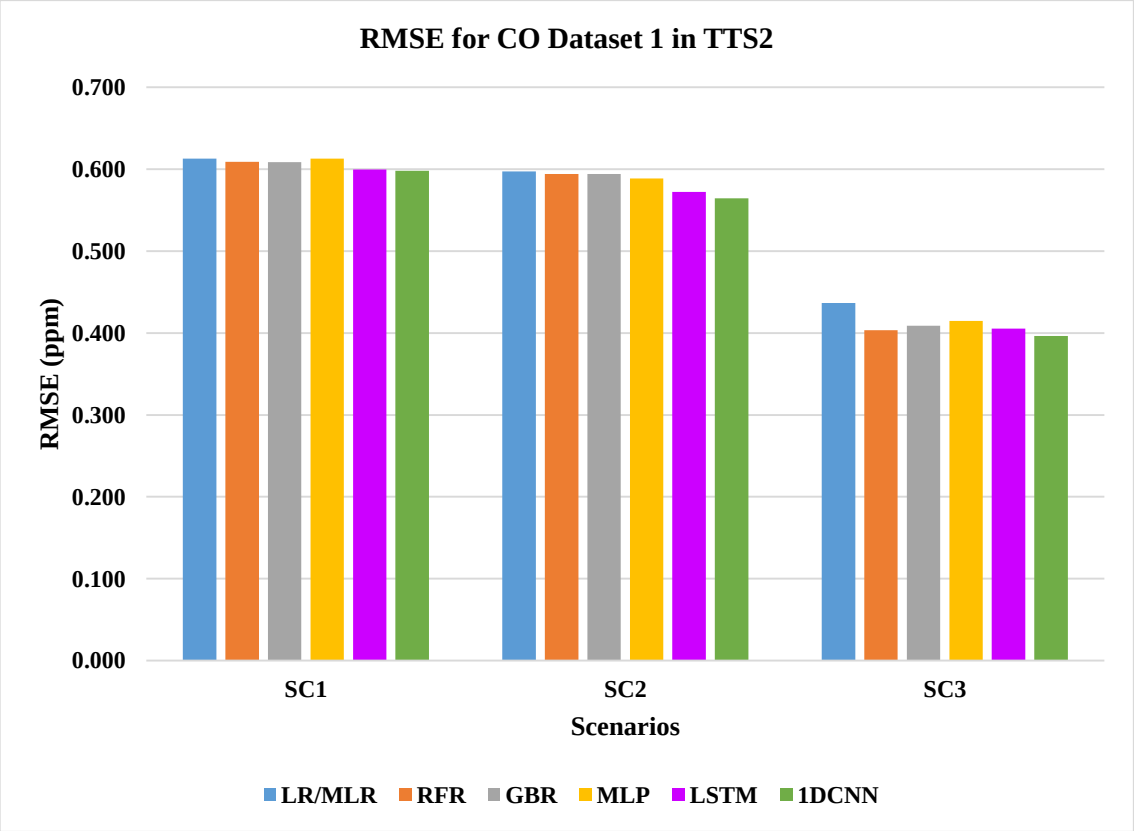


Figure 4.2: RMSE comparison between different scenarios and algorithms for CO Dataset 1 in TTS2.

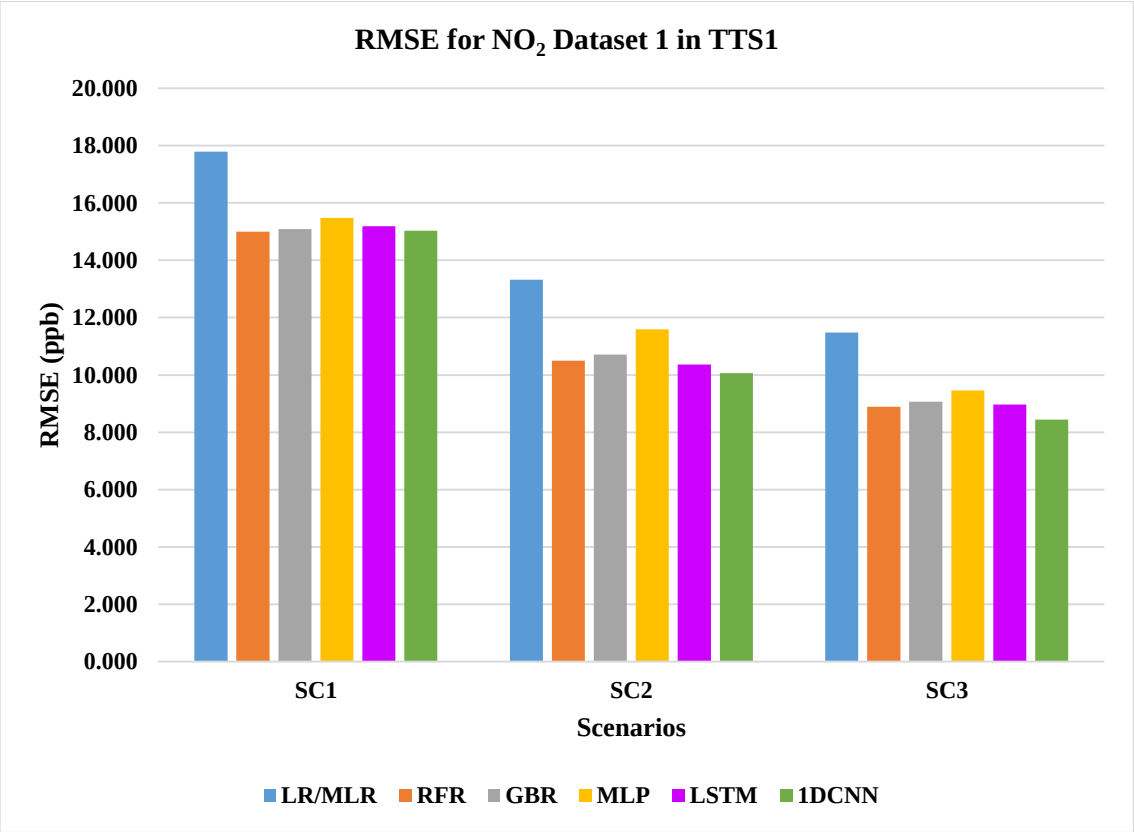


Figure 4.3: RMSE comparison between different scenarios and algorithms for NO_2 Dataset 1 in TTS1.

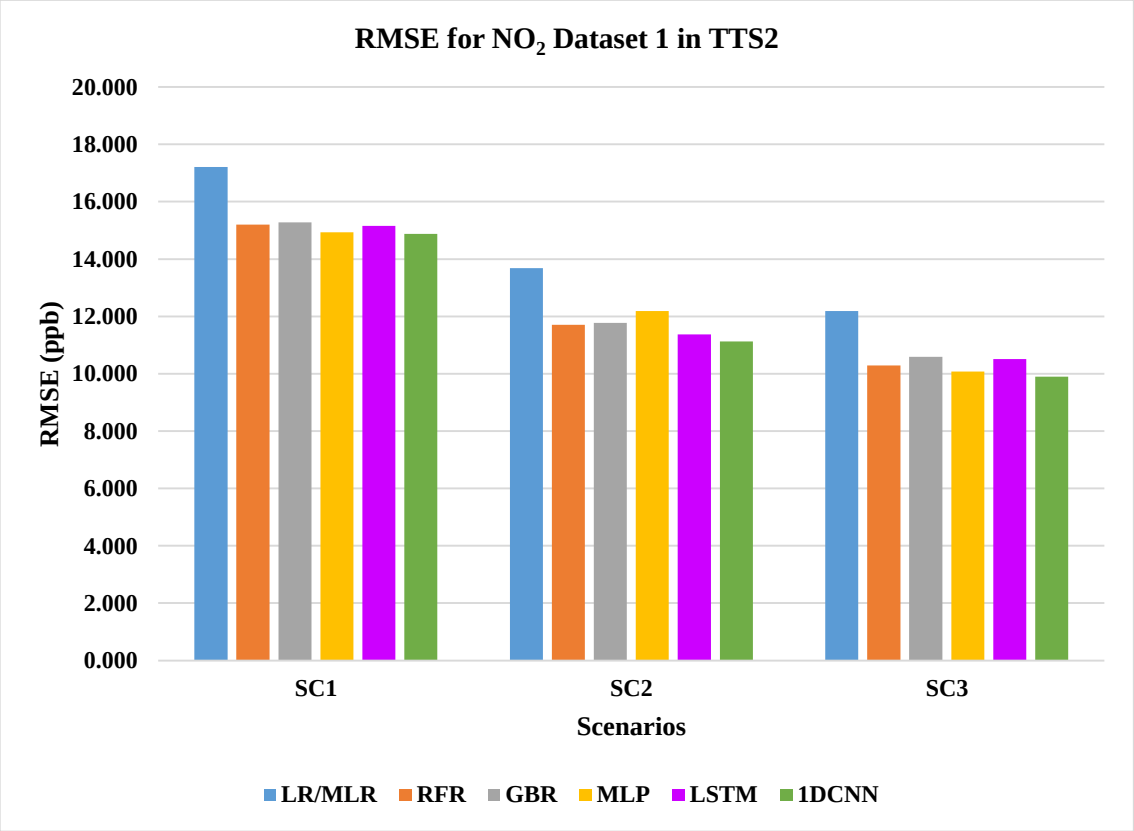


Figure 4.4: RMSE comparison between different scenarios and algorithms for NO_2 Dataset 1 in TTS2.

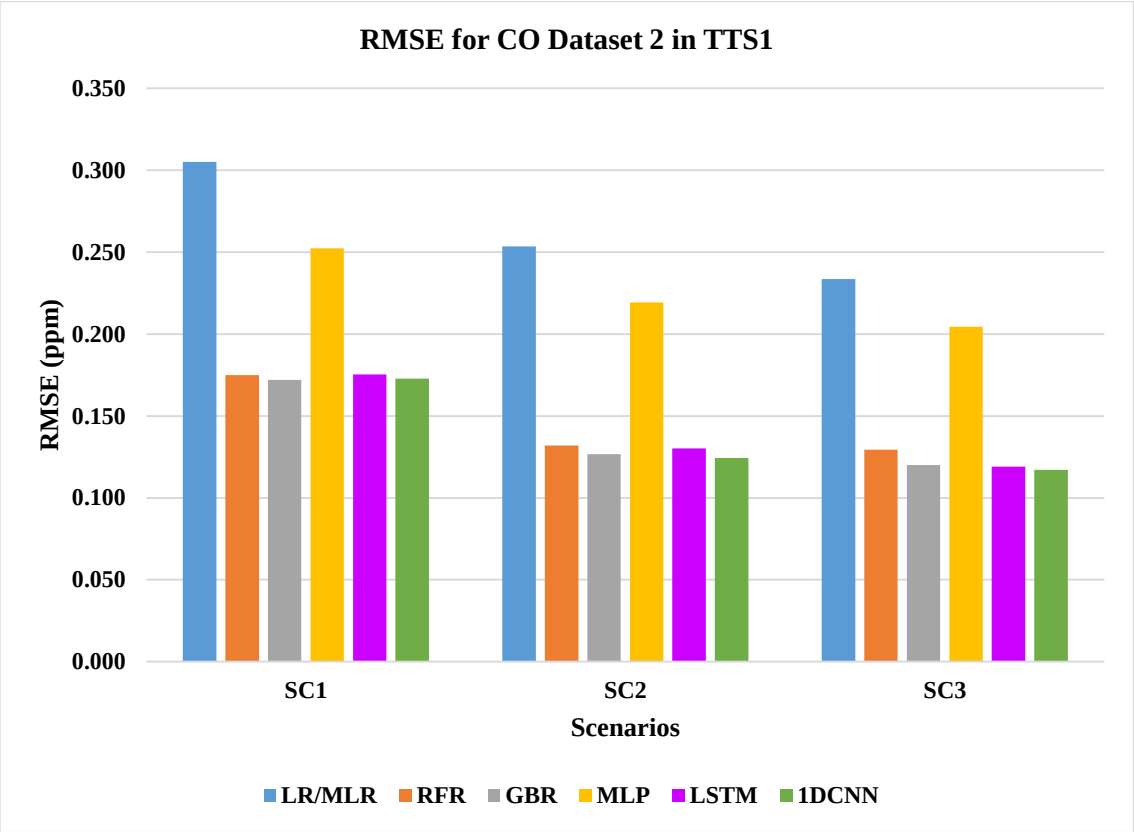


Figure 4.5: RMSE comparison between different scenarios and algorithms for CO Dataset 2 in TTS1.

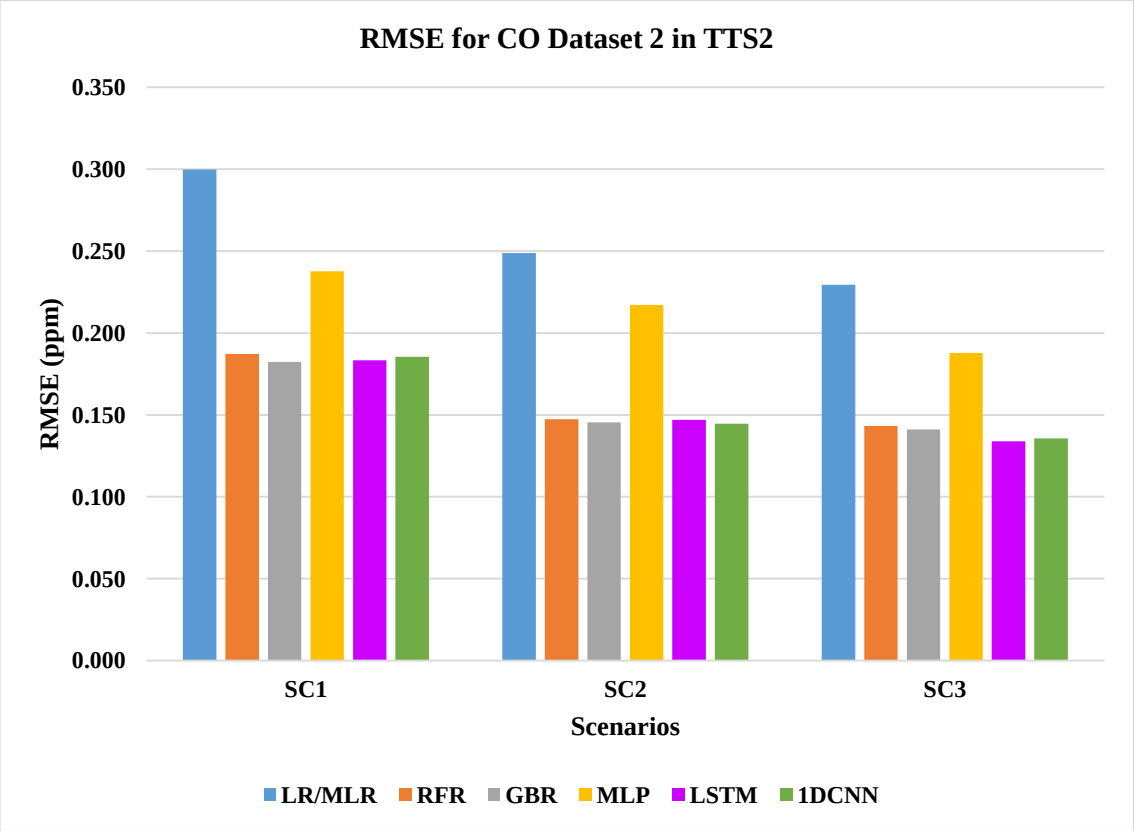


Figure 4.6: RMSE comparison between different scenarios and algorithms for CO Dataset 2 in TTS2.

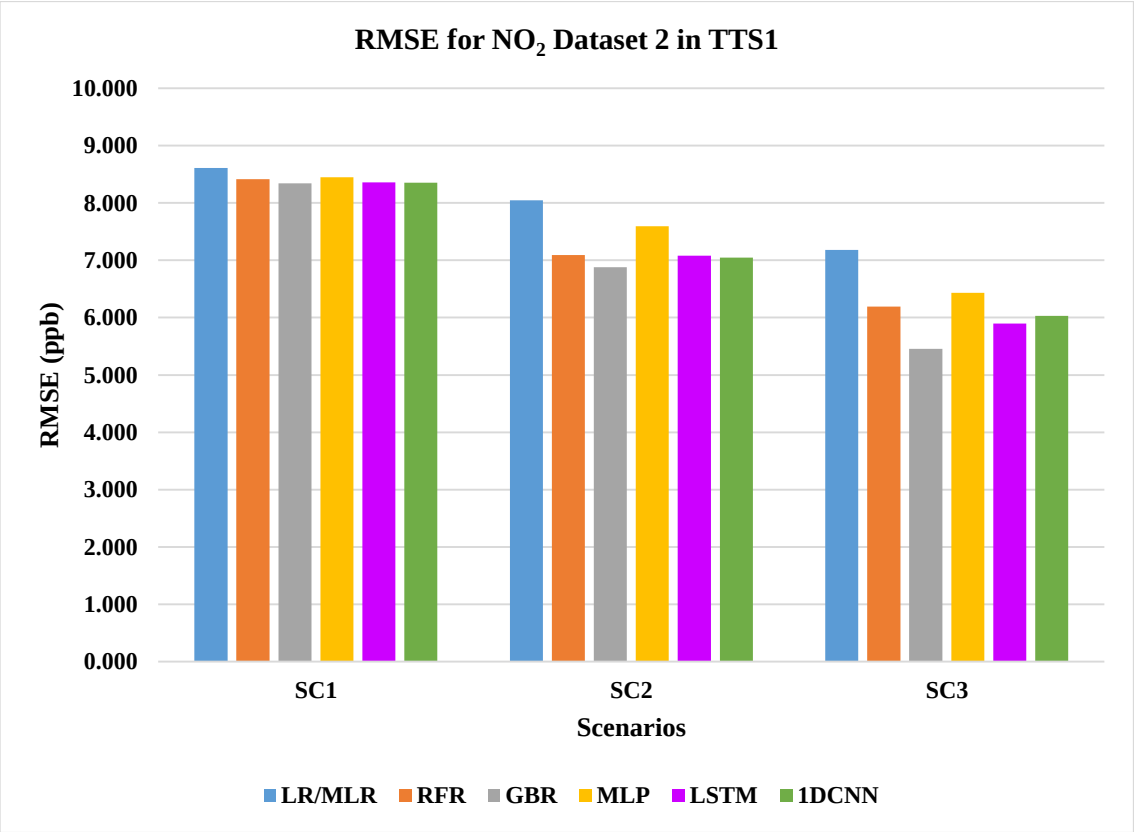


Figure 4.7: RMSE comparison between different scenarios and algorithms for NO_2 Dataset 2 in TTS1.

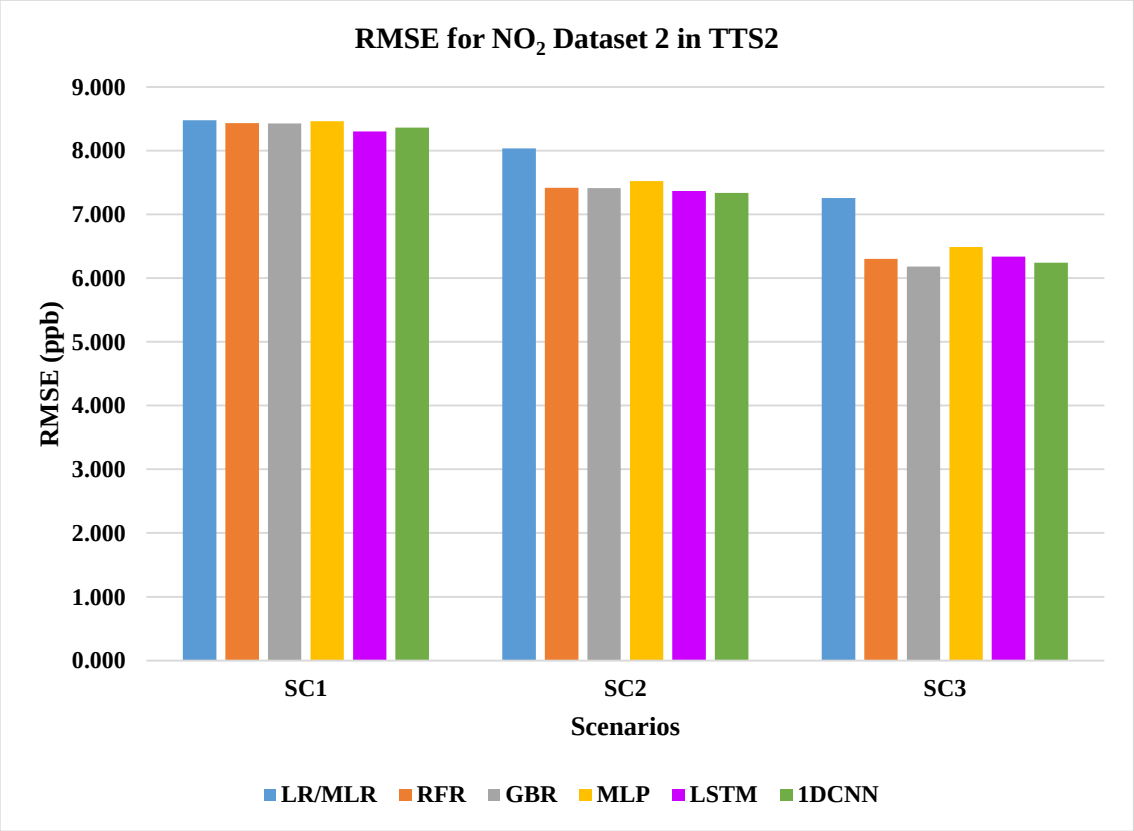


Figure 4.8: RMSE comparison between different scenarios and algorithms for NO_2 Dataset 2 in TTS2.

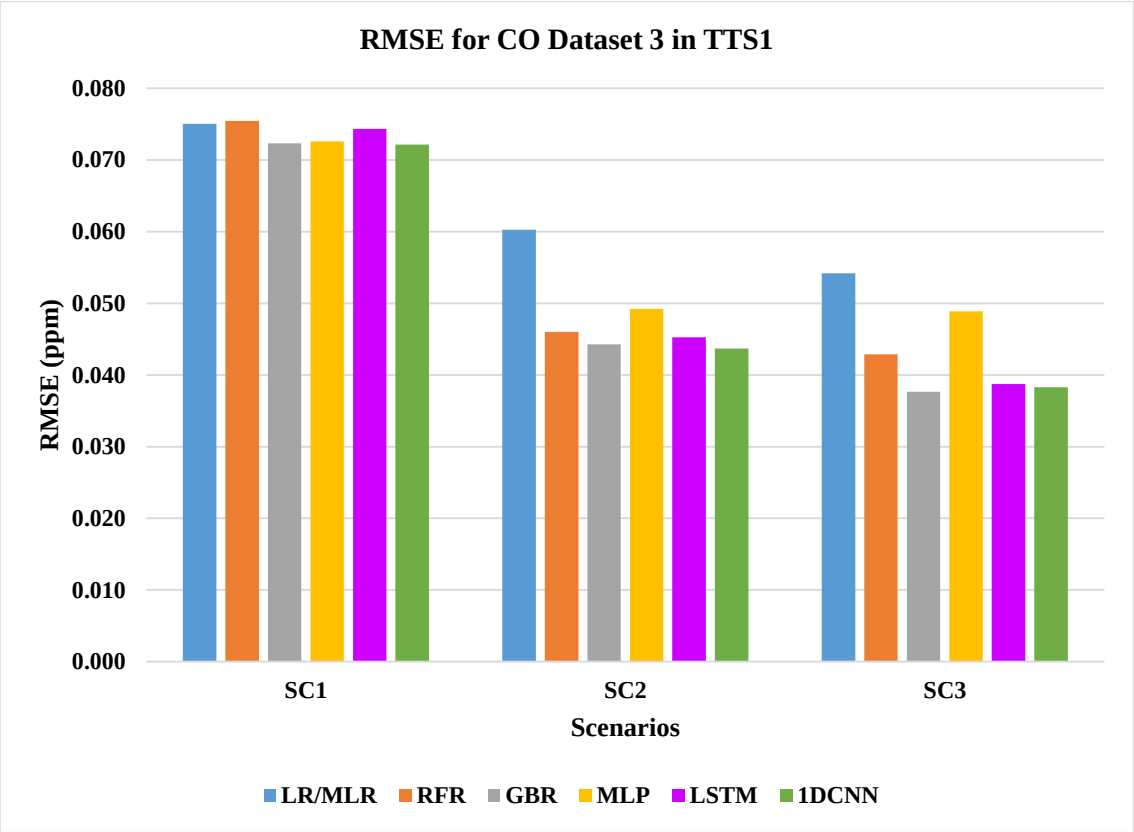


Figure 4.9: RMSE comparison between different scenarios and algorithms for CO Dataset 3 in TTS1.

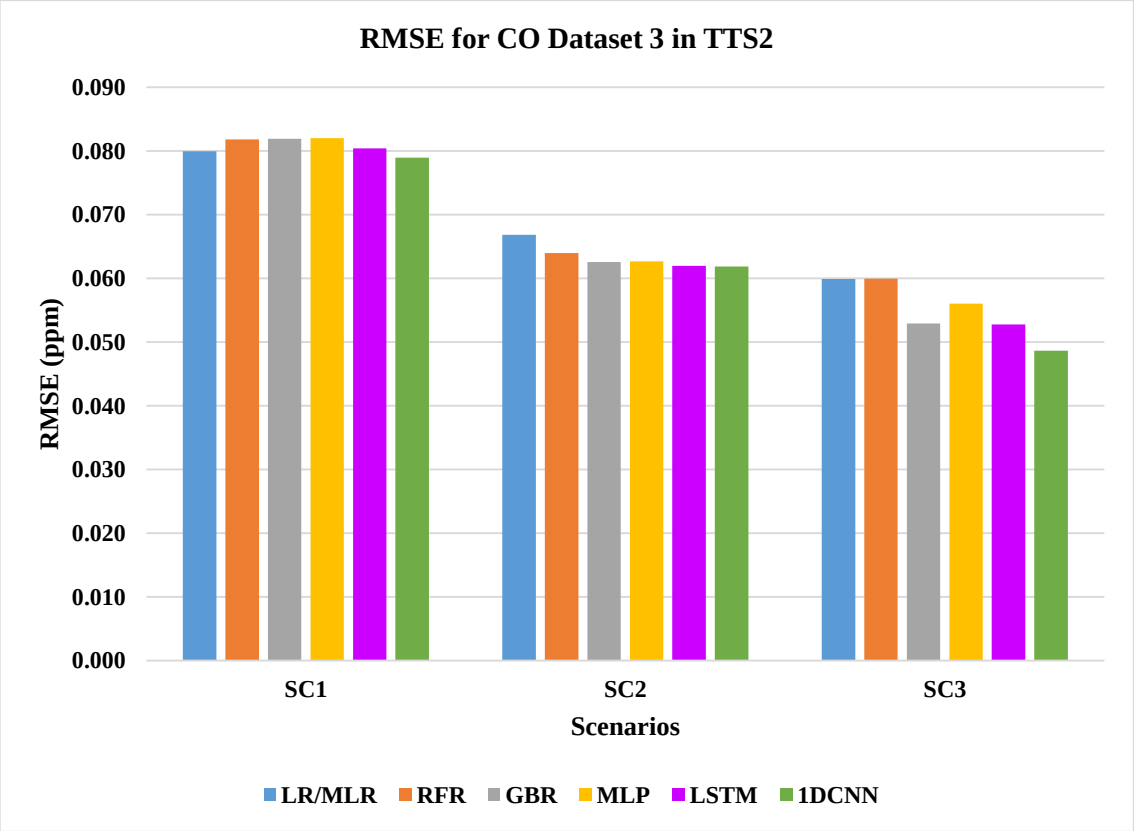


Figure 4.10: RMSE comparison between different scenarios and algorithms for CO Dataset 3 in TTS2.

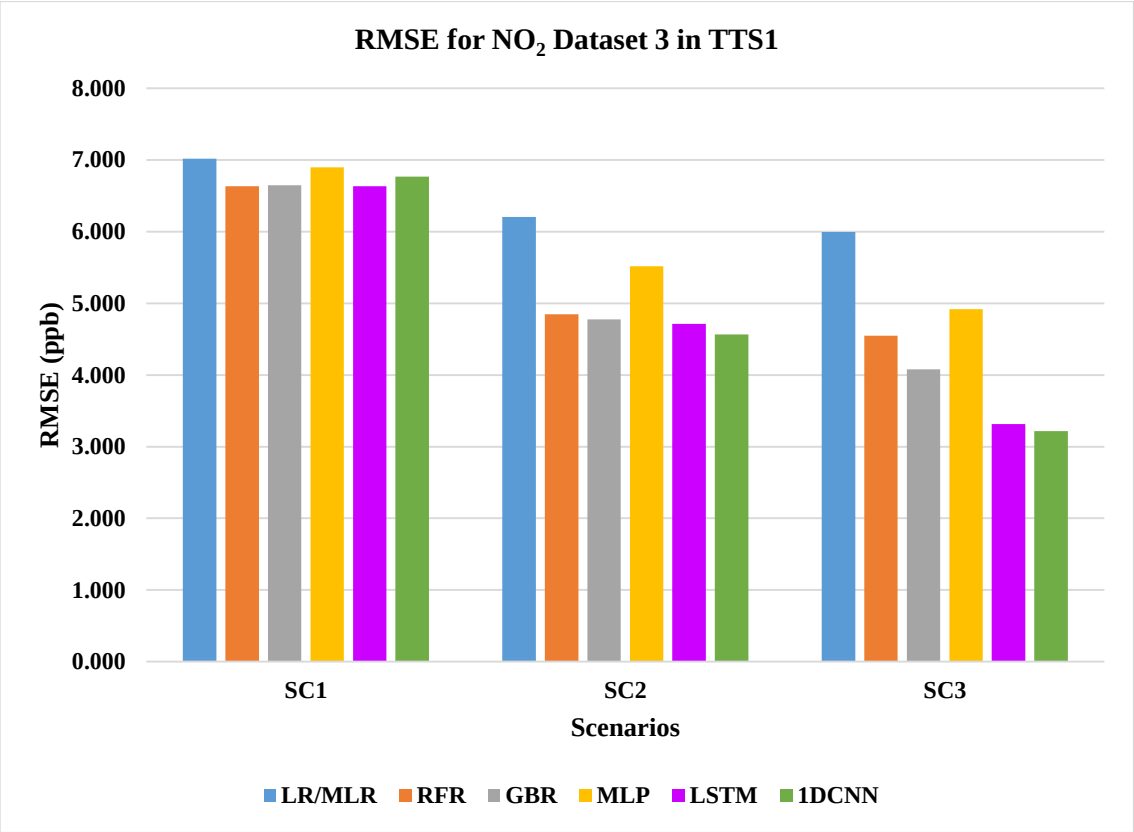


Figure 4.11: RMSE comparison between different scenarios and algorithms for NO_2 Dataset 3 in TTS1.

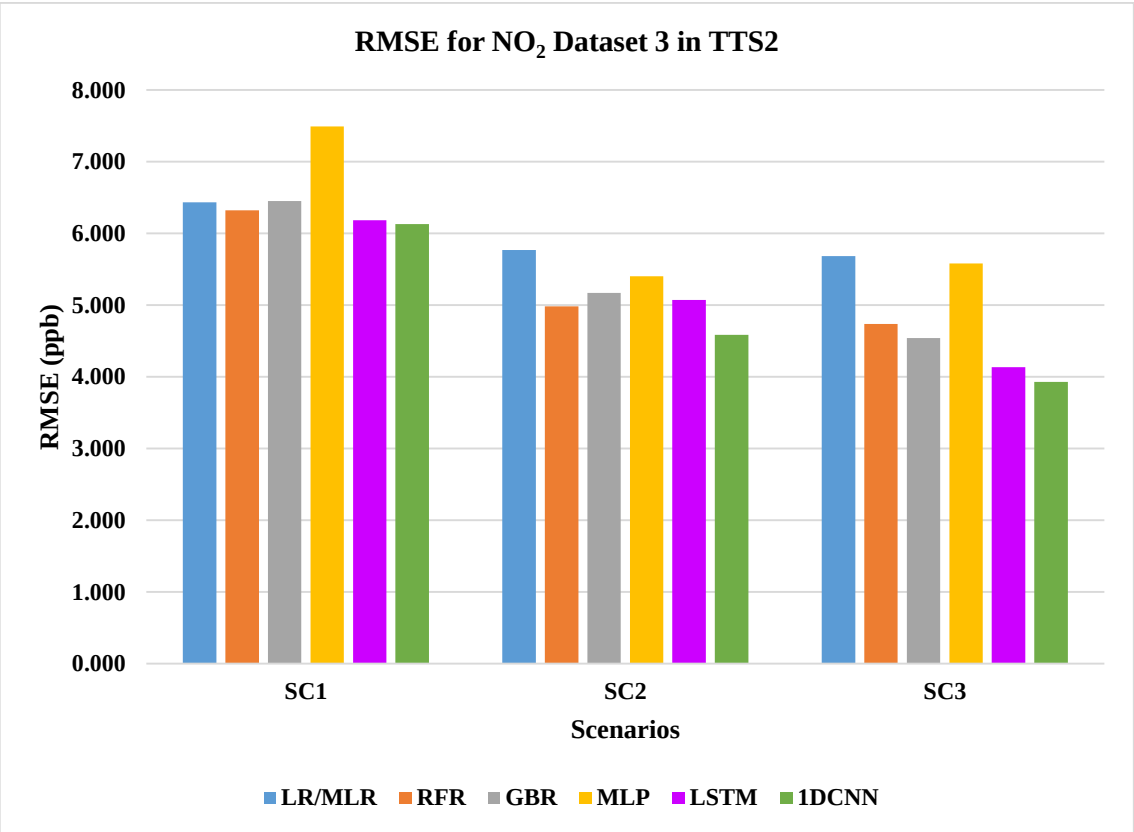


Figure 4.12: RMSE comparison between different scenarios and algorithms for NO_2 Dataset 3 in TTS2.

4.4 Summary of Results

The accuracy of every calibration algorithm improves (lower RMSE, higher R^2) from SC1 to SC3. The accuracy improves when the temperature (T) and humidity (RH) are included (SC2) alongside the raw target pollutant (CO or NO_2) data. Similarly, better accuracy can be observed when other pollutant information is introduced to exploit the dependencies arising from cross-sensitivity. However, the scale of improvement in SC2 and SC3 from the SC1 scenario differs across the datasets and across the two pollutant (CO and NO_2) gases.

For the CO gas calibration, the accuracy improvement in SC3 is far more prominent than the improvement in SC2 from SC1 in Dataset 1. While the improvements of accuracy in SC3 from SC1 do not show a similar jump in Dataset 2 and 3. However, the improvement in SC2 from SC1 is more noteworthy for both Datasets 2 and 3 than that of Dataset 1. This clearly shows that the impact of cross-sensitive gases on LCS response is more than temperature and relative humidity in Dataset 1 than the other two datasets for CO. Similarly, it can be said that the impact of temperature and relative humidity on the LCS data was more in Datasets 2 and 3 than Dataset 1. On the other hand, for NO_2, SC2 is noted to get a significant improvement from SC1 in Dataset 1 and then a very modest improvement in SC3, showing the effect of temperature and relative humidity on the LCS response was more than cross-sensitive gases (quite the opposite of what has been observed for CO in Dataset 1).

For both pollutants (CO and NO_2), there is a clear additional improvement in accuracy when all other pollutant raw data are included alongside temperature (T) and humidity (RH) (SC3). This improvement clearly emphasizes the importance of deploying LCS as multi-sensor platforms. Not only that allows for monitoring multiple pollutants with a single unit, but the accuracy of the measured data also improves through mitigation of cross-sensitivity.

All algorithms have outperformed the linear regression-based calibration methods for all scenarios. In almost every instance, 1DCNN is the best-performing algorithm. These results show that 1DCNN could significantly improve the accuracy of low-cost multi-sensor air pollutant monitors. LSTM, GBR, and RFR are the next most accurate algorithms, and their performances are quite close in many instances. While LSTM and RFR have gained much traction for gas sensor calibration, *1DCNN and GBR-based*

calibrations appear to have received far less attention and therefore warrant strong consideration from researchers.

4.4.1 Effects of different TTS and Scenarios

The accuracy of any given algorithm is better for the 90/10 split (TTS1) compared to the 20/80 split (TTS2). Interestingly, the accuracy improvement from going from SC1 to SC3 and from SC2 to SC3 is more noticeable than going from TTS2 to TTS1. The co-variate factors seem to have more impact than longer training/co-location time for CO. For example, consider the RMSE of the top three performing algorithms for all three datasets in CO calibration (Tables 4.7, 4.9 and 4.11). Overall, the RMSE improves more when going from SC1 to SC3 than from TTS2 to TTS1 in all cases. Similarly, the covariate factors also appear more important than longer training/co-location time for NO_2. The RMSE improvement from SC1 to SC3 for TTS1 and improvement from TTS2 to TTS1 for SC1 of the top three performing algorithms in CO and NO_2 calibration for all three datasets are illustrated in Table 4.13 and 4.14 as examples. The following observations can be made from these two tables:

1. In CO calibration (Table 4.13), RMSE improvements in TTS1 from TTS2 (for SC1) in all three datasets are mostly less than 10%, whereas RMSE improvements in SC3 from SC1 (for TTS1) are over 30% for Dataset 1 and 2 and more than 40% in Dataset 3.
2. Similarly, in NO_2 calibration (Table 4.14), the RMSE improvements in SC3 scenario in TTS1 is significant (30% - 50%). While introducing more training data for SC1 has little to no impact (when compared between TTS1 and TTS2).
3. However, above perspective changes for both CO and NO_2 calibrations in the RMSE improvements from TTS2 to TTS1 for SC3 scenario. The RMSE improvements in TTS1 from TTS2 for SC3 are several times better than that of SC1. Interestingly, the RMSE improvements in SC3 from SC1 (for TTS1) are still higher than them.

Table 4.13: RMSE improvement in CO for the top three performing algorithms.

Dataset	Algorithm	Improvement from SC1 to SC3 (for TTS1)	Improvement from TTS2 to TTS1 (for SC1)	Improvement from TTS2 to TTS1 (for SC3)
Dataset 1	1DCNN	36.54%	9.53%	13.34%
	LSTM	36.92%	9.17%	15.17%
	RFR	37.64%	9.03%	14.33%
Dataset 2	1DCNN	32.16%	6.49%	13.66%
	LSTM	32.12%	4.37%	11.12%
	GBR	30.30%	5.49%	15.01%
Dataset 3	1DCNN	46.95%	8.86%	21.31%
	LSTM	47.92%	7.50%	26.65%
	GBR	47.90%	12.20%	28.82%

Table 4.14: RMSE improvement in NO_2 for the top three performing algorithms.

Dataset	Algorithm	Improvement from SC1 to SC3 (TTS1)	Improvement from TTS2 to TTS1 (for SC1)	Improvement from TTS2 to TTS1 (for SC3)
Dataset 1	1DCNN	43.79%	-1.01%	14.65%
	GBR	39.88%	1.28%	14.39%
	RFR	40.74%	1.35%	13.65%
Dataset 2	1DCNN	27.79%	0.10%	3.39%
	LSTM	29.45%	-0.67%	7.01%
	GBR	34.64%	0.99%	11.82%
Dataset 3	1DCNN	52.44%	-1.04%	18.09%

	LSTM	50.01%	-7.34	19.76%
	GBR	38.60%	-2.96%	10.16%

The impact on RMSE improvement for having temperature, relative humidity, and other cross pollutant gas data as input parameters is quite vital. The accuracy of the models derived and evaluated with TTS2 is not significantly worse than those for TTS1. This seems to suggest that with sophisticated calibration models like the ones presented in this chapter, not only the LCS platforms could be utilized as a backup for a reference grade monitor (TTS1), but they can also be deployed for reasonably accurate CO and NO_2 monitoring for a long duration after a short co-location (TTS2). It should be noted that the accuracy of the calibration models could be further improved by further periodic co-location and recalibration (please see [16]).

4.4.2 Empirical CDFs

Figure 4.13 shows the empirical CDF (E-CDF) of absolute errors for 1DCNN calibration techniques (other figures for the top three calibration techniques are added in the Appendix). In all cases, SC3 has demonstrated better performance than SC2, where both cross-sensitivity and T and RH correction are dealt with.

The previously mentioned comparison of improvement in RMSE values for all scenarios from TTS2 to TTS1 can also be observed in the E-CDFs (all E-CDF figures can be found in the Appendix). Figure 4.13 is an example of that comparison for 1DCNN. From these figures also, it can be established that the co-variate factors are quite important, and it is not just about ensuring longer training/co-location time.

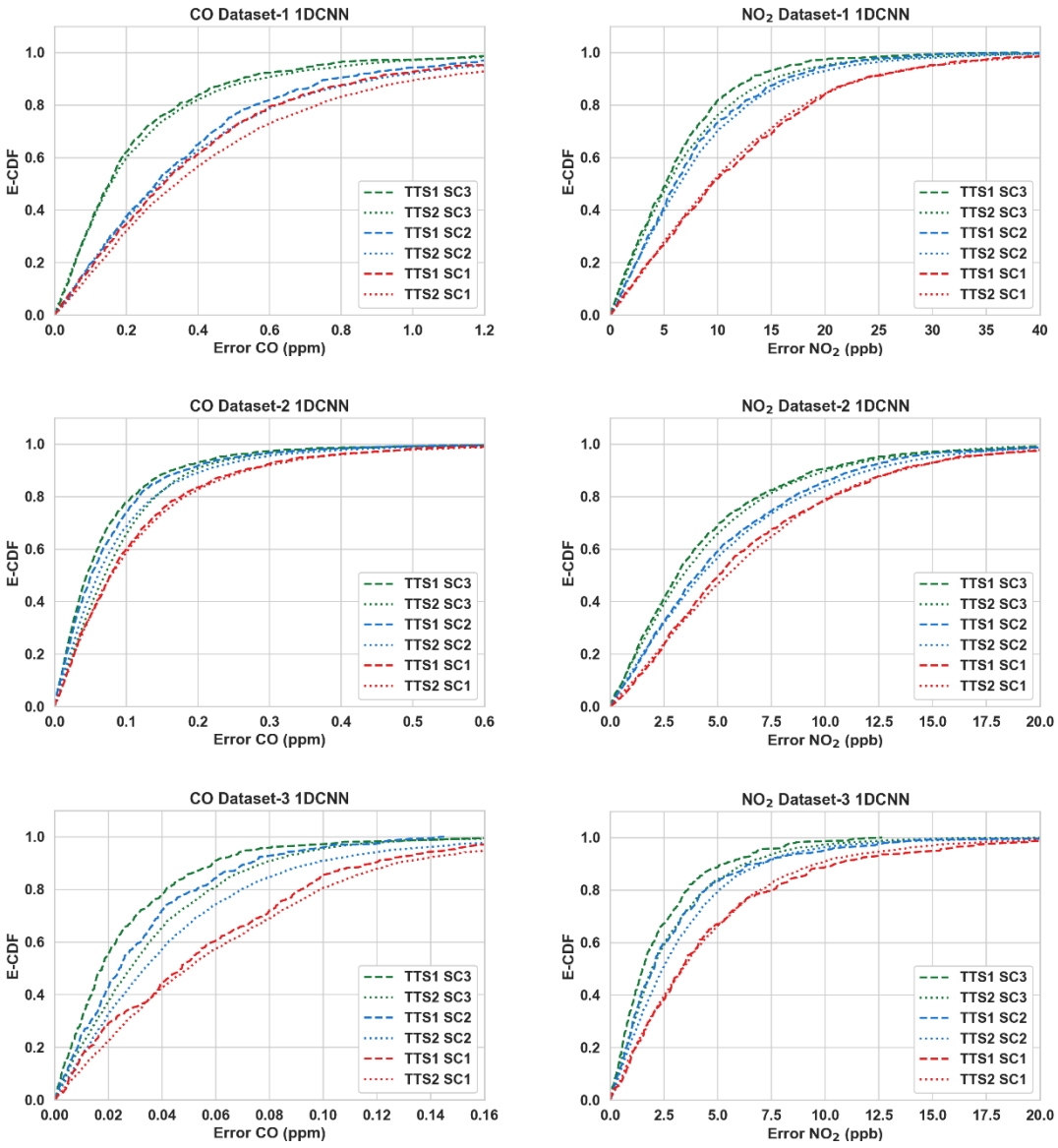


Figure 4.13: Empirical CDFs of absolute errors for 1DCNN calibration.

4.4.3 Target Diagrams

Target diagrams for the used calibration models have been observed, and the target diagrams for 1DCNN are shown in Figure 4.14 and 4.15 as examples (please see all target diagrams in the Appendix). The following observations are made from the target diagrams:

1. All points lie within the unit circle (radius = 1); therefore, the variance of the residuals is smaller than the variance of the reference measurements. It is an essential characteristic of a functional calibration model [31], indicating that the variability of the dependent variable (calibrated output) is explained by the

independent variable (the reference data) and not the residual [53]. It should be noted that all calibration algorithms presented in this work fulfil this criterion.

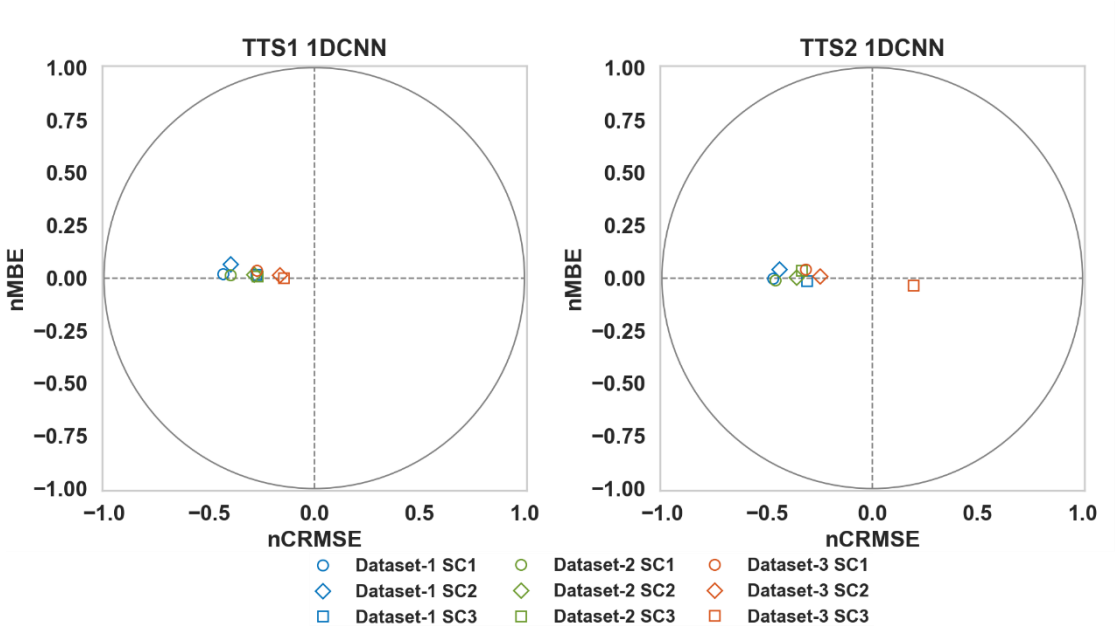


Figure 4.14: Target diagrams of 1DCNN algorithm for CO.

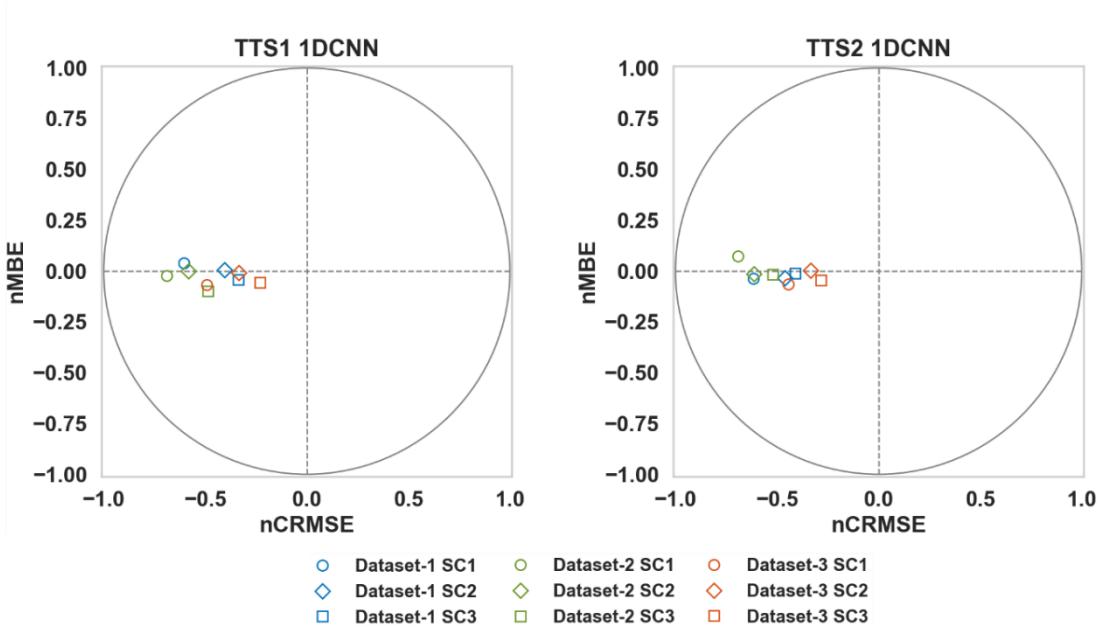


Figure 4.15: Target diagrams of 1DCNN algorithm for NO_2.

2. The distance from the origin, which measures the normalized RMSE (RMSE/$\sigma_{reference}$) clearly show that the SC3 regressors are more accurate than the other scenario regressors. This accuracy once again demonstrates the importance of the availability of co-variate factors like temperature, relative

humidity, other gas pollutants, and time variables. The normalized CRMSE (nCRMSE), normalized MBE (nMBE), and normalized RMSE (nRMSE) values for the algorithms from all three datasets are attached in the Appendix.

3. Most of the points lie on the left, indicating that the standard deviation of the calibrated sensor data for most models is smaller than the ground truth standard deviation.
4. For TTS1, the points lie above the x-axis, indicating that the models, on average, slightly overestimate both gas concentrations. For TTS2, a few models also slightly underestimate both gas concentrations.

CHAPTER 5

RESULTS & DISCUSSIONS: CO-VARIATES

This chapter proposes the utilization of new input parameters or co-variates to improve the calibration performance. These input parameters were feature engineered from the readily available timestamp data that came with all LCS and reference sensor data. This unique approach had a positive impact on the performance of the algorithms and such method has never been thoroughly explored in the literature to the best of our knowledge.

5.1 Introduction

As discussed earlier, the response of the LCS are highly susceptible to cross-sensitivity from other ambient gases [14, 57] and temperature and relative humidity [14, 53]. Therefore, temperature, relative humidity and cross pollutant data are used as regressor co-variates to calibrate the LCS output. Chapter 4 presents a thorough investigation of the impact of these parameters on various calibration algorithms.

It is well known that the performance of the LCS drifts and degrade over time [10, 103, 167]. We hypothesize that the number of days a sensor is in operation can be used as a co-variate enabling the ML algorithms to model and address the gradual degradation. Many gas pollutants are directly produced from human activities (e.g., CO, NO_2 resulting from automobile emissions) [12, 17]. Therefore, it is reasonable to assume that the time (or hour) of the day that influences the relevant human activities will also impact the pollutant concentration and should be used as a co-variate factor. However, the literature does not show evidence of utilizing these factors, which are readily available without any additional cost, for multi-variate calibration of LCS. Our investigation, as discussed in this chapter, showed that including these parameters as input, features can significantly improve the accuracy of the sensor calibration.

5.2 Feature Engineering

Feature engineering is the process of selecting, manipulating, and transforming raw samples or existing features into new features that can be used to make supervised machine learning perform better [223, 224]. The datasets used in this work are time series datasets; all the samples come with additional information of the time they were recorded. However, this feature has yet to be fully used in any of the previous studies of LCS calibration. In this work, the information of time presented in the datasets was transformed to multiple new features and added as new features for the supervised machine learning techniques used to calibrate the gas sensors. The new features generated from the timestamped information are the hour of the timestamp ($Hour$) and number of days deployed (N_{day}). The exact values of the hour were extracted from the already existing timestamped information and number of days deployed (N_{day}) was calculated based on the length of days the sensor was deployed to collect data.

5.3 Experimental Scenarios

Based on the summary from the previous section, we have looked into four more scenarios (S1, S2, S3 and S4) to understand the effect of the number of days in operation and time (hour) of the day values on the calibration accuracy.

5.3.1 Scenario 1 (S1)

This involves deriving regressors or calibration models so that

$$P_{calibrated}^{S1} = \Phi^{S1}\{P_{raw}, T, RH, GAS_{raw}\} \quad 5.1$$

The regressor, Φ^{S1}, is derived solely based on the raw pollutant sensor input (working electrode data and/or auxiliary electrode data) along with temperature and relative humidity readings to minimize the MSE between $P_{calibrated}^{S1}$ and the ground truth. This is the identical to SC3 from Chapter 4. For the ease of further comparison, we have renamed it to S1 for this section.

5.3.2 Scenario 2 (S2)

The second case uses the number of days the sensor has been in deployment, N_{day} , as another input so that

$$P_{calibrated}^{S2} = \Phi^{S2}\left\{P_{raw}, T, RH, GAS_{raw}, N_{day}\right\} \quad 5.2$$

The regressor, Φ^{S2}, is now using number of days of deployment as an input variable in addition to the inputs for S1 to minimize the MSE between $P_{calibrated}^{S2}$ and the ground truth.

5.3.3 Scenario 3 (S3)

Scenario 3 replaces N_{day} with time of the day ($Hour$) as input while keeping the other inputs (temperature, relative humidity, and cross pollutant) the same as Scenario 2. Therefore, the regressor, Φ^{S3}, is now using the time of the day that each sample was collected along with other inputs to minimize the MSE between $P_{calibrated}^{S3}$ and the ground truth.

$$P_{calibrated}^{S3} = \Phi^{S3}\left\{P_{raw}, T, RH, GAS_{raw}, Hour\right\} \quad 5.3$$

5.3.4 Scenario 4 (S4)

Both N_{day} and $Hour$ are included as input parameters in scenario 4 along with the other inputs. Therefore, the regressor, Φ^{S4} can be written as

$$P_{calibrated}^{S4} = \Phi^{S4}\left\{P_{raw}, T, RH, GAS_{raw}, N_{day}, Hour\right\} \quad 5.4$$

We have used four algorithms for each dataset from the previous chapter to investigate the effects of time-related variables N_{day} and $Hour$. 1DCNN has been selected due to its consistently excellent performance. LSTM and GBR algorithms represent RNN and ensemble ML techniques respectively. MLR, the most commonly reported algorithm, was employed to benchmark the performance of the ML-based methods. The same rigorous training, validation, and testing method, outlined in Chapter 3, has been followed during this work, where the hyperparameters have been tuned on the relevant training datasets and tested on the corresponding testing sets for the regressors. The following results represent the 90/10 split of the training and testing data (TTS1).

5.4 Results and Observations

Table 5.1 – 5.6 illustrates the performance improvements in the above mentioned scenarios for the three best overall performing algorithms from last chapter, namely, 1DCNN, LSTM and GBR, and benchmarked them with respect to MLR. The following observations can be made from these results:

1. Overall, the use of N_{day} and $Hour$ has improved the calibration accuracy in both pollutants for all three datasets significantly. In all cases, S4 showed the best RMSE scores.
2. For both pollutants (CO and NO_2), MLR has failed to capture the essence of the time variables and did not show any noticeable changes to the RMSE or broadly saying on the algorithm's accuracy. MLR showed some impact in Dataset 3 of NO_2 gas, compared to other results, especially in S2 and S4.
3. Other algorithms have shown a large improvement in accuracy regarding lower RMSE values. Especially 1DCNN and GBR have considerably improved scores in S2 and S4 scenarios from S1 for both CO and NO_2.

Table 5.1: Performance improvement of CO in dataset 1 for TTS1.

Algorithm	Scenario	S1	S2	S3	S4
1DCNN	RMSE (ppm)	0.344	0.298	0.314	0.264
	Improvement of RMSE from S1 (in %)		13.17	8.53	23.05
LSTM	RMSE (ppm)	0.344	0.335	0.326	0.322
	Improvement of RMSE from S1 (in %)	0	2.63	5.17	6.44
GBR	RMSE (ppm)	0.349	0.306	0.333	0.295
	Improvement of RMSE from S1 (in %)		12.20	4.59	15.33
MLR	RMSE (ppm)	0.38	0.37	0.39	0.37
	Improvement of RMSE from S1 (in %)		3.03	-0.42	2.78

Table 5.2: Performance improvement of CO in dataset 2 for TTS1.

Algorithm	Scenario	S1	S2	S3	S4
1DCNN	RMSE (ppm)	0.117	0.105	0.111	0.093
	Improvement of RMSE from S1 (in %)		10.59	5.14	20.29
LSTM	RMSE (ppm)	0.119	0.110	0.117	0.109
	Improvement of RMSE from S1 (in %)	0	7.54	1.58	8.66
GBR	RMSE (ppm)	0.120	0.106	0.110	0.093
	Improvement of RMSE from S1 (in %)		11.22	7.96	22.18
MLR	RMSE (ppm)	0.234	0.227	0.233	0.226
	Improvement of RMSE from S1 (in %)		2.68	0.10	3.02

Table 5.3: Performance improvement of CO in dataset 3 for TTS1.

Algorithm	Scenario	S1	S2	S3	S4
1DCNN	RMSE (ppm)	0.038	0.026	0.038	0.026
	Improvement of RMSE from S1 (in %)		32.78	1.50	32.86
LSTM	RMSE (ppm)	0.039	0.029	0.038	0.027
	Improvement of RMSE from S1 (in %)	0	24.91	2.97	30.89
GBR	RMSE (ppm)	0.038	0.025	0.038	0.025
	Improvement of RMSE from S1 (in %)		33.79	0.29	32.66
MLR	RMSE (ppm)	0.054	0.054	0.054	0.054
	Improvement of RMSE from S1 (in %)		0.53	0.10	0.60

Table 5.4: Performance improvement of NO_2 in dataset 1 for TTS1.

Algorithm	Scenario	S1	S2	S3	S4
1DCNN	RMSE (ppb)	8.445	7.355	8.400	6.888
	Improvement of RMSE from S1 (in %)		12.91	0.53	18.44
LSTM	RMSE (ppb)	8.968	8.560	8.836	8.476
	Improvement of RMSE from S1 (in %)	0	4.55	1.47	5.49
GBR	RMSE (ppb)	9.069	7.368	8.472	6.871
	Improvement of RMSE from S1 (in %)		18.76	6.58	24.24
MLR	RMSE (ppb)	11.486	11.496	11.338	11.354
	Improvement of RMSE from S1 (in %)		-0.09	1.28	1.15

Table 5.5: Performance improvement of NO_2 in dataset 2 for TTS1.

Algorithm	Scenario	S1	S2	S3	S4
1DCNN	RMSE (ppb)	6.030	5.136	5.114	4.535
	Improvement of RMSE from S1 (in %)		14.83	15.20	24.80
LSTM	RMSE (ppb)	5.896	5.603	5.736	5.342
	Improvement of RMSE from S1 (in %)	0	4.97	2.71	9.39
GBR	RMSE (ppb)	5.453	5.066	4.938	4.374
	Improvement of RMSE from S1 (in %)		7.09	9.45	19.79
MLR	RMSE (ppb)	7.183	7.187	7.184	7.190
	Improvement of RMSE from S1 (in %)		-0.06	-0.02	-0.09

Table 5.6: Performance improvement of NO_2 in dataset 3 for TTS1.

Algorithm	Scenario	S1	S2	S3	S4
1DCNN	RMSE (ppb)	3.219	3.116	3.200	3.141
	Improvement of RMSE from S1 (in %)		3.18	0.56	2.41
LSTM	RMSE (ppb)	3.317	3.084	3.207	2.958
	Improvement of RMSE from S1 (in %)	0	7.02	3.32	10.80
GBR	RMSE (ppb)	4.079	3.383	3.896	3.408
	Improvement of RMSE from S1 (in %)		17.08	4.50	16.46
MLR	RMSE (ppb)	5.995	5.546	5.989	5.537
	Improvement of RMSE from S1 (in %)		7.48	0.09	7.64

4. Generally, more improvement has been observed in S2 than S3 throughout all these experiments. The effects of the use of N_{day} is more impactful than $Hour$ as an additional input parameter. A combination of both in S4 has illustrated a further improved RMSE. The improvement of accuracy in S4 is clearly not a linear addition of the improvements seen in S2 and S3. Thus, a sophisticated ML algorithm (e.g., 1DCNN) is much needed to calculate this improved accuracy in S4.
5. For CO, the maximum improvements can be seen in dataset 3 in S2 and S4 scenarios, where the improvements are over 30% for 1DCNN and GBR. LSTM also showed a noticeable improvement, close to 25% in S2 and just over 30% in S4 for dataset 3. The RMSE improvements in S2 and S4 are around 10% and 20% respectively for both 1DCNN and GBR in Dataset 1 and 2 for CO. Improvement for LSTM in these two datasets is slightly lower than 1DCNN and GBR.
6. For NO_2, both 1DCNN and GBR showed better improvement scores in S2 and S4 than other algorithms. The percentage of improvement was around 20% in Datasets 1 and 2 for 1DCNN and GBR. RMSE improvement in S3 was

noticeably bigger in Dataset 2 for NO_2 than in other scenarios. Also, Dataset 3 showed the least improvement of the three datasets for NO_2. In Dataset 3, RMSE improvements (for NO_2) were mostly below 10% in most cases, except GBR for S2 and S4.

The previous results show the RMSE scores for TTS1 where the training datasets were larger. Similar experiments were conducted for TTS2 (20/80 split of the training and testing data) with a smaller training dataset. Tables 5.7 and 5.8 show the improvement from S1 to S4 (using both variables N_{day} and $Hour$) for TTS2 for CO and NO_2 respectively. TTS2 also showed noticeable accuracy improvements for 1DCNN, GBR and LSTM in S4 for both gases. However, the improvements in TTS2 are more modest than its TTS1 counterpart. A smaller training dataset, meaning less temporal information, is the likely reason behind this decrease. Regardless, for both cases (TTS1 and TTS2), the incorporation of N_{day} and $Hour$ have demonstrably improved the accuracy of the ML-based calibration algorithms. The empirical CDF plots of calibration error for 1DCNN in Figure 5.1 show a clear improvement in S4 from S1 in both TTS1 and TTS2, further underscoring the importance of using both N_{day} and $Hour$ data as input features. Similar to the observations made in Chapter 4, the algorithms in TTS1 have performed better than those in TTS2. Also, the best-performing scenario is the TTS1 and S4 in all cases for both gases.

Table 5.7: Performance improvement of CO for TTS2.

Algorithm	**Improvement of CO RMSE in S4 from S1 (in %)**		
	Dataset 1	**Dataset 2**	**Dataset 3**
1DCNN	7.83	13.97	12.24
LSTM	6.34	4.13	19.24
GBR	5.87	15.60	5.66
MLR	2.97	2.18	0.14

Table 5.8: Performance improvement of NO_2 for TTS2.

Algorithm	Improvement of NO_2 RMSE in S4 from S1 (in %)		
	Dataset 1	Dataset 2	Dataset 3
1DCNN	13.26	13.10	8.55
LSTM	4.49	3.55	2.36
GBR	18.24	13.50	9.05
MLR	1.39	-0.04	6.79

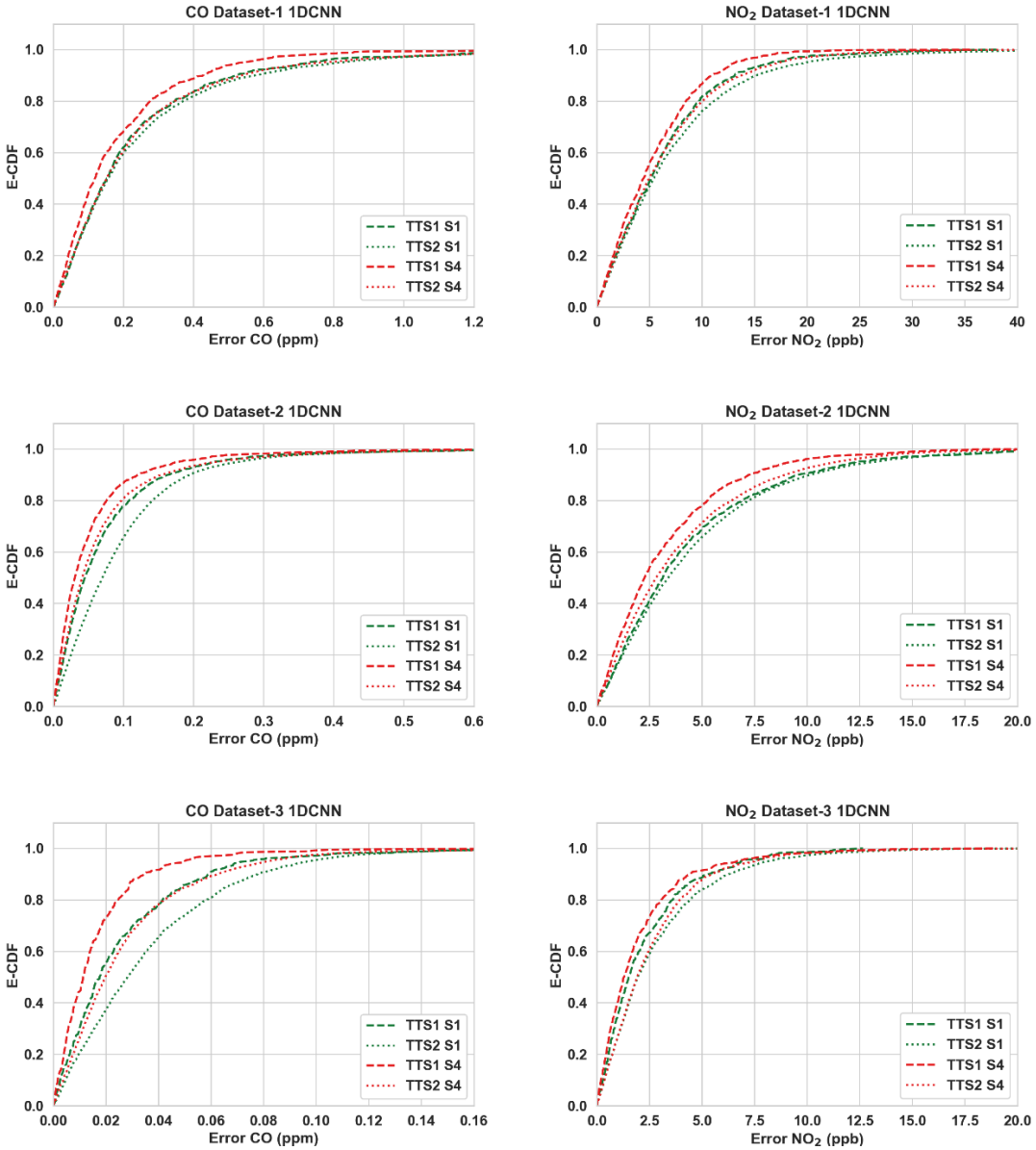
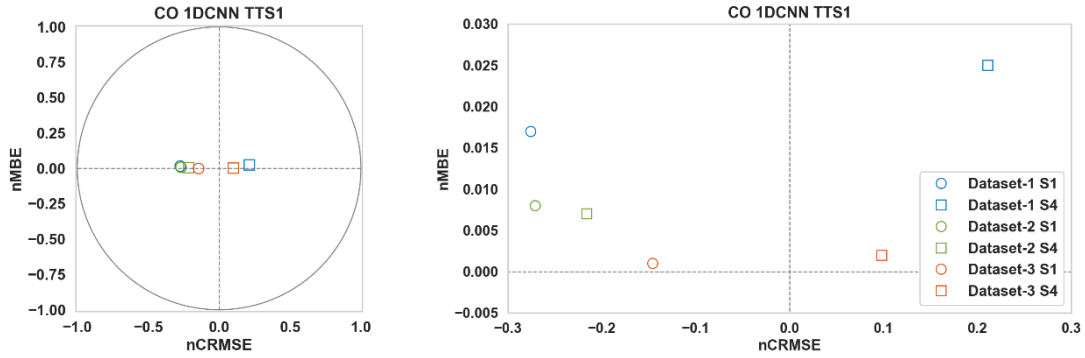
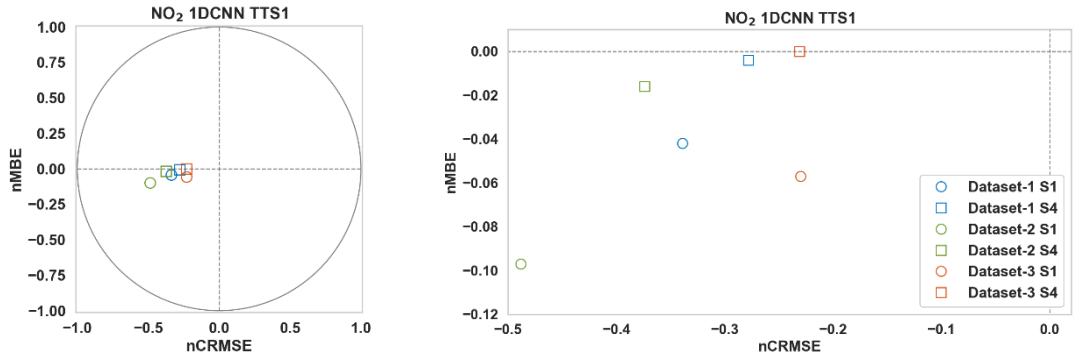


Figure 5.1: Empirical CDF plots of calibration error for 1DCNN for all datasets.

The target diagrams for the 1DCNN calibration are presented in Figure 5.2. All the points lie inside the unit circle (radius = 1). Thus, the variance of the residuals is smaller than that of the reference measurements, and the variability of the calibrated output (dependent variable) is explained by the reference data (independent variable) and not the residual. The distance of these points from the origin, i.e., the normalised RMSE show that S4 calibrations are more accurate than S1. This again proves the importance of adding N_{day} and $Hour$ data as input features. It is also observed that the standard deviation of the calibrated data is mostly smaller than the standard deviation of the ground truth as the majority of the points lie on the left plane. Also, it can be seen that the points for CO reside on the positive y-axis while they are on the negative side of the y-axis for NO_2. It means the results for CO 1DCNN are slightly overestimated while all the NO_2 results for 1DCNN are slightly underestimated.



(a)



(b)

Figure 5.2: Target diagrams for the 1DCNN calibration of (a) CO and (b) NO_2 for all datasets.

5.5 Observations

Many low-cost air quality sensors measure only one target pollutant. Temperature and relative humidity sensors are usually on board with the pollutant sensor as these two ambient sensors are cheap. Thus, additional information on the temperature and relative humidity are gathered at a low-cost along with the pollutant data. However, in many field deployments, several gas pollutant sensors are put together in the form of a sensor array to acquire multiple pollutant concentrations. For example, all three datasets used in this study utilized array of sensors to measure other ambient pollutants along with the target gas, temperature, and relative humidity. There are two main advantages of assembling a sensor array like this.

Firstly, multiple pollutant gases can be measured using the same setup or deployment thus, broadening the sensor array's utility.

Also, low-cost gas sensors are susceptible to cross-sensitivity to other ambient gases. Having additional data on other pollutant gases can significantly improve calibration accuracy. This has been well investigated in the literature and systematically verified in the results described in the previous chapter.

However, these benefits come with the additional cost of multiple sensors. Based on the efficacy of the temporal co-variates established in this chapter, we propose a cost-effective alternative: utilizing the number of days deployed (N_{day}) and time of day ($Hour$) data as input for the calibration algorithms. Consider the case of utilizing these two readily available co-variates alongside T and RH (termed SC2T). We have used similar methodology to train and validate the algorithms. Results for 1DCNN and GBR, the two most consistent algorithms found in this study and not utilized in the literature, are shown in Table 5.9. All the results show a noticeable improvement in RMSE after calibration. It is obvious that the accuracy of the calibration can be significantly improved even without deploying a sensor array of multiple pollutants and therefore without increasing the cost.

The accuracy of the calibration improves on both occasions, SC3 and SC2T. We have compared these results between SC3 and SC2T. For this, the improvement of RMSE scores in SC3 and SC2T from SC2 for TTS1 are calculated and shown in Table 5.10 and 5.11. Overall, the improvements for SC2T are substantial and not far behind the

improvements observed in SC3 for most cases and for a few instances, SC2T seems on par or even slightly outperform SC3.

Table 5.9: Improvement of RMSE in SC2T (raw + temperature + humidity + N_{day} + $Hour$) from SC2 (raw + temperature + humidity) for 1DCNN and GBR for TTS1.

Target Pollutant	Algorithm	Improvement of RMSE in SC2T from SC2 (in %)		
		Dataset 1	Dataset 2	Dataset 3
CO	1DCNN	16.57	3.96	16.38
	GBR	16.99	4.21	12.01
NO_2	1DCNN	1.91	6.16	25.16
	GBR	5.55	2.99	16.54

Table 5.10: Improvement of RMSE in SC3 (raw + temperature + humidity + other gases) and SC2T (raw + temperature + humidity + N_{day} + $Hour$) from SC2 (raw + temperature + humidity) for CO in TTS1.

Algorithm	Scenario	Improvement of CO RMSE in SC3 or SC2T from SC2 (in %) for TTS1		
		Dataset 1	Dataset 2	Dataset 3
MLR	SC3	29.20	7.88	10.02
	SC2T	10.50	6.00	10.06
RFR	SC3	31.89	1.83	6.78
	SC2T	15.73	0.91	10.36
GBR	SC3	31.03	5.34	14.93
	SC2T	16.99	4.21	12.01
MLP	SC3	33.38	6.77	0.77
	SC2T	15.24	18.30	0.70
LSTM	SC3	31.41	8.58	14.48
	SC2T	15.53	2.74	11.95
1DCNN	SC3	31.58	5.77	12.39
	SC2T	16.57	3.96	16.38

Table 5.11: Improvement of RMSE in SC3 (raw + temperature + humidity + other gases) and SC2T (raw + temperature + humidity + N_{day} + $Hour$) from SC2 (raw + temperature + humidity) for NO_2 in TTS1.

Algorithm	Scenario	Improvement of NO_2 RMSE in SC3 and SC2T from SC2 (in %) for TTS1		
		Dataset 1	Dataset 2	Dataset 3
MLR	SC3	13.81	10.71	3.40
	SC2T	3.07	1.35	5.32
RFR	SC3	15.38	12.63	6.18
	SC2T	3.12	5.36	10.08
GBR	SC3	15.31	20.73	14.58
	SC2T	5.55	3.00	16.53
MLP	SC3	18.36	15.31	10.89
	SC2T	7.04	9.71	10.09
LSTM	SC3	13.49	16.73	29.59
	SC2T	1.85	5.21	20.89
1DCNN	SC3	16.04	14.40	29.53
	SC2T	1.91	6.16	25.15

Table 5.12 and 5.13 show the similar results from 1DCNN and GBR for TTS2; the improvement pattern presented here is similar to TTS1 from the previous tables.

Table 5.12: Improvement of RMSE in SC3 (raw + temperature + humidity + other gases) and SC2T (raw + temperature + humidity + N_{day} + $Hour$) from SC2 (raw + temperature + humidity) for CO in TTS2.

Algorithm	Scenario	Improvement of CO RMSE in SC3 and SC2T from SC2 (in %) for TTS2		
		Dataset 1	Dataset 2	Dataset 3
1DCNN	SC3	29.74	6.15	21.35
	SC2T	17.15	2.30	24.69
GBR	SC3	31.19	2.97	15.45
	SC2T	16.44	1.55	17.85

Table 5.13: Improvement of RMSE in SC3 (raw + temperature + humidity + other gases) and SC2T (raw + temperature + humidity + N_{day} + $Hour$) from SC2 (raw + temperature + humidity) for NO_2 in TTS2.

Algorithm	**Scenario**	**Improvement of NO_2 RMSE in SC3 and SC2T from SC2 (in %) for TTS2**		
		Dataset 1	**Dataset 2**	**Dataset 3**
1DCNN	**SC3**	11.05	14.94	14.23
	SC2T	1.99	5.25	11.56
GBR	**SC3**	10.04	16.56	12.14
	SC2T	4.65	5.68	15.25

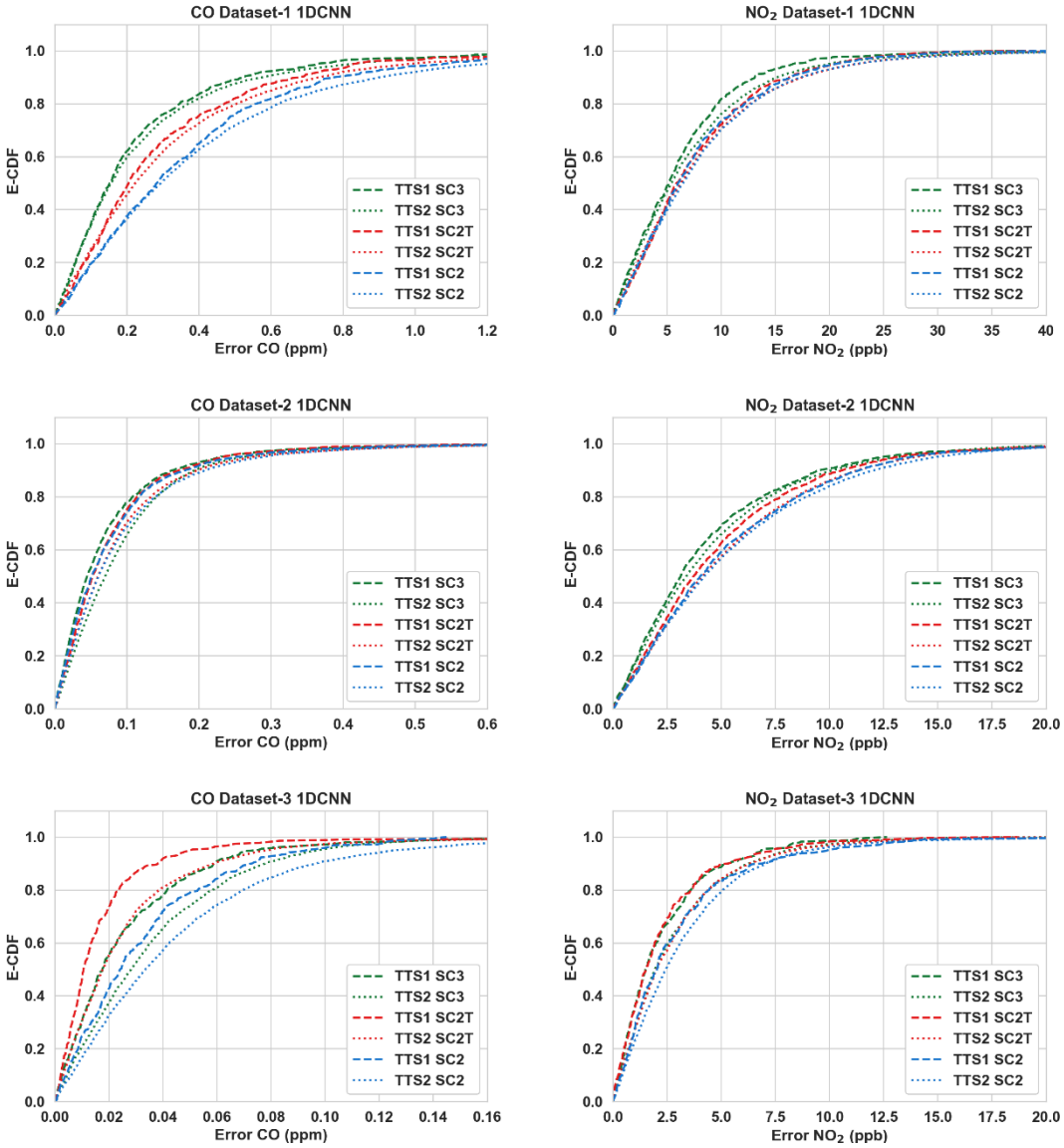
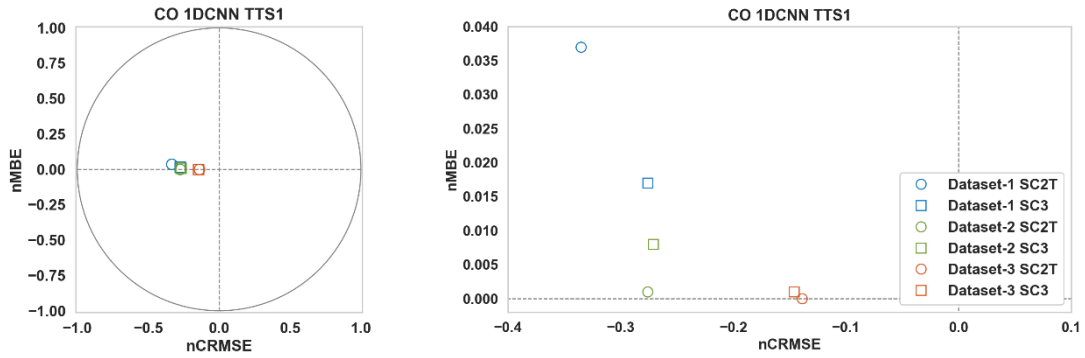
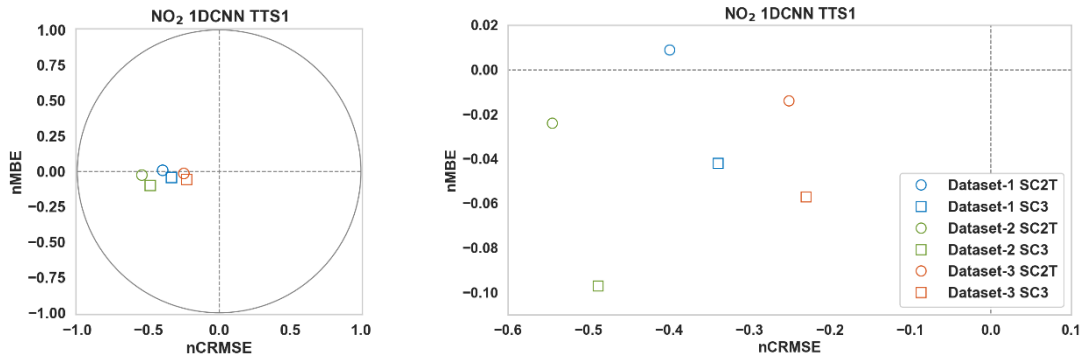


Figure 5.3: Empirical CDF plots for 1DCNN comparing SC2 (raw + temperature + humidity), SC3 (raw + temperature + humidity + other gases) and SC2T (raw + temperature + humidity + N_{day} + $Hour$) errors for both TTS1 and TTS2 for all datasets.

The empirical CDF plots are illustrated in Figure 5.3 for 1DCNN. It can be observed that the SC2T performance is quite close to SC3 generally. The target diagrams of the SC3 and SC2T presented in Figure 5.4 for 1DCNN show similar improvement patterns.



(a)



(b)

Figure 5.4: Target diagrams comparing SC3 (raw + temperature + humidity + other gases) and SC2T (raw + temperature + humidity + N_{day} + $Hour$) for 1DCNN calibration of (a) CO and (b) NO_2 in TTS1 for all datasets.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

This research work proposes and develops novel machine learning-based algorithms for the calibration of low-cost ambient gas pollutant sensors. A robust calibration methodology was established to study and benchmark the proposed algorithms with their counterparts reported in the literature. The two most important findings of this work were:

1) The importance of leveraging the temporal information, e.g., time of deployment and time of day. This information is readily available and has so far has not been utilized in the reported literature for Low-Cost Sensors (LCS) calibration.

2) The potential benefit of using One Dimensional Convolutional Neural Network (1DCNN) and Gradient Boosting Regression (GBR) for the calibration of low-cost ambient gas pollutant sensors. Both the algorithms have so far been largely overlooked in the literature and in our investigation show consistently effective performance across a variety of scenarios.

The investigation conducted in this thesis has directly resulted in the publication of the following journal articles:

i. **S. Ali**, F. Alam, K. Arif, and J. Potgieter, "Leveraging Temporal Information to Improve Machine Learning-Based Calibration Techniques for Low-Cost Air Quality Sensors", *Sensors*, May 2024. (**Q1, IF 3.9**)

ii. **S. Ali**, F. Alam, K. Arif, and J. Potgieter, "Low-Cost CO Sensor Calibration Using One Dimensional Convolutional Neural Network", *Sensors*, January 2023. (**Q1, IF 3.9**)

iii. **S. Ali**, T. Glass, B. Parr, J. Potgieter and F. Alam, "Low Cost Sensor with IoT LoRaWAN Connectivity and Machine Learning Based Calibration for Air Pollution Monitoring," *IEEE Transactions on Instrumentation & Measurement*, October 2020. (**Q1, IF 5.6**)

The presented results and discussions can be summarized in the following key points:

1. Three field deployment scenarios are found in the reported literature for LCS calibration. First scenario (SC1) only considers the raw LCS data that targets the pollutant gas measurement. Second scenario (SC2) includes temperature and relative humidity along with the raw LCS data. The third scenario (SC3) adds other cross-sensitive gas data collected by a LCS sensor array with the second scenario. Out of these three scenarios found in the literature, SC3 is typically preferred (sensor array) for calibration as this leads to higher accuracy. However, depending on the hardware utilized in the sensors, this improvement in SC3 from SC2 may not always be substantial. another point that should be kept in mind while using a sensor array for SC3 is that utilizing a sensor array will increase the overall cost and complexity of the LCS. However, this will also significantly improve the potential usability of the LCS as the sensor array can now be used to measure multiple pollutants.

2. Development and evaluation of machine learning-based calibration methods require training (and validation) and testing sets. In this thesis, two different train-test splitting ratios were used, namely TTS1 and TTS2. TTS1 used 90% of the entire dataset as training and 10% for testing, whereas TTS2 used a much smaller training set (20% of the dataset) and a larger testing set (80% of the dataset). From this study it was found that having more deployment data to train the algorithms was always better than having less. The performance for TTS1 was more accurate than that for TTS2 for every algorithm. Based on this, we can surmise that a co-located low-cost sensor can potentially be used as a backup in case the reference grade monitor is out of commission for a short period due to fault or maintenance However, TTS2 results were still reasonably when the algorithms were appropriately trained. Therefore, it might be possible to deploy LCS for ambient pollutant monitoring after co-locating with a reference sensor for a set period for calibration.

3. 1DCNN-based multi-variate regressors are proposed for the calibration of LCS. This calibration algorithm was benchmarked against other ML-based calibration techniques and linear regression and was found to be the most accurate. GBR-based calibration models also, in general, perform better than many of the popular

ML techniques for all three datasets. Both of these algorithms have not been utilized and explored in the literature and can potentially improve the performance of low-cost sensors that monitor ambient gas pollutants.

Using temporal parameters as input for the algorithms improved the performance of the calibration. It can be concluded that using temporal variables like time in deployment (N_{day}) and time of the day ($Hour$) can improve calibration accuracy significantly. This is noteworthy since this can be achieved without increasing the cost. These temporal data are available regardless. As discussed earlier, incorporating other cross-sensitive gas sensors in the LCS can always achieve more accuracy. Even in such a case, the calibration accuracy can be improved further by including the readily available information in the form of the temporal variables as input co-variates.

6.2 Future Work

The following works can be pursued to explore this research work further:

1. The continual progress in deep learning and big data analysis present the opportunity to use new and advanced ML algorithms that have not been used in this research.
2. It is not clear how the ML models behave when the trained model from one LCS is used to calibrate another LCS with similar hardware and configuration. It will be worthwhile to investigate how a transfer calibration approach can be used to migrate such scenario. The idea would be to train one calibration model using the data from one co-located LCS and then test the accuracy of the model with the data collected by another LCS. The calibrated model can be a *baseline* that will then be further trained (transfer calibration) to work well for another LCS. Unfortunately, the three datasets used in this thesis came from the field deployment of three different sensor arrays and therefore did not allow us to test the efficacy of transfer calibration. In order to perform this study, multiple LCS with similar hardware needs be deployed together. Transfer calibration approach has not been investigated very well in the literature and needs be studied in future.
3. Our investigation with the three datasets showed that the time of deployment and time of the day have significant impact on the accuracy of the algorithms when

used as input parameters. However, there are other available temporal parameters such as number of years in deployment, month of the year, date of the month, whether the day is a weekday or weekend etc. The impact of these parameters may vary depending on the length of the deployment, and some may have more influence than others due to a change in pollution patterns caused by human and industrial activities at different times and locations. While these mentioned temporal parameters were found to have no noticeable impact for the three datasets in this work, a future study with other datasets may show them to be useful co-variates.

4. A suite of LCS can be developed and deployed for further data collection to validate/verify the findings. Several low-cost sensor arrays were constructed at the very initial stage of this study with this aim. Three CO sensor nodes were deployed with a co-located reference sensor in Auckland, New Zealand for a short period [222]. However, due to the COVID-19 pandemic and lockdowns, no further investigations were possible. A fresh attempt can be taken with the findings and guidelines from this study to further investigate real-time and long-term LCS deployments.

REFERENCES

[1] W. H. Organization, "Ambient air pollution: A global assessment of exposure and burden of disease," 2016.

[2] A. Goshua, C. A. Akdis, and K. C. Nadeau, "World Health Organization global air quality guideline recommendations: Executive summary," *Allergy,* vol. 77, no. 7, pp. 1955-1960, 2022.

[3] A. J. Cohen *et al.*, "Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015," *The Lancet,* vol. 389, no. 10082, pp. 1907-1918, 2017.

[4] M. Renzi *et al.*, "A nationwide study of air pollution from particulate matter and daily hospitalizations for respiratory diseases in Italy," *Science of the Total Environment,* vol. 807, p. 151034, 2022.

[5] W. Nazar and M. Niedoszytko, "Air pollution in Poland: a 2022 narrative review with focus on respiratory diseases," *International journal of environmental research and public health,* vol. 19, no. 2, p. 895, 2022.

[6] C. Nishida and K. Yatera, "The impact of ambient environmental and occupational pollution on respiratory diseases," *International Journal of Environmental Research and Public Health,* vol. 19, no. 5, p. 2788, 2022.

[7] J. Colls and A. Tiwary, *Air pollution: Measurement, modelling and mitigation,* 3rd Edition ed. Taylor & Francis Group, 2017, p. 528.

[8] K. B. Shaban, A. Kadri, and E. Rezk, "Urban Air Pollution Monitoring System With Forecasting Models," *IEEE Sensors Journal,* vol. 16, no. 8, pp. 2598-2606, 2016, doi: 10.1109/JSEN.2016.2514378.

[9] A. C. Rai *et al.*, "End-user perspective of low-cost sensors for outdoor air pollution monitoring," *Science of The Total Environment,* vol. 607, pp. 691-705, 2017.

[10] S. De Vito *et al.*, "Crowdsensing IoT Architecture for Pervasive Air Quality and Exposome Monitoring: Design, Development, Calibration, and Long-Term Validation," *Sensors,* vol. 21, no. 15, p. 5219, 2021.

[11] N. Castell *et al.*, "Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?," *Environment international,* vol. 99, pp. 293-302, 2017.

[12] K. Hu, V. Sivaraman, B. G. Luxan, and A. Rahman, "Design and Evaluation of a Metropolitan Air Pollution Sensing System," *IEEE Sensors Journal,* vol. 16, no. 5, pp. 1448-1459, 2016, doi: 10.1109/JSEN.2015.2499308.

[13] P. Kumar *et al.*, "The rise of low-cost sensing for managing air pollution in cities," *Environment international,* vol. 75, pp. 199-205, 2015.

[14] Y. Liang *et al.*, "Field comparison of electrochemical gas sensor data correction algorithms for ambient air measurements," *Sensors and Actuators B: Chemical,* vol. 327, p. 128897, 2021/01/15/ 2021, doi: https://doi.org/10.1016/j.snb.2020.128897.

[15] J. M. Barcelo-Ordinas, J. Garcia-Vidal, M. Doudou, S. Rodrigo-Muñoz, and A. Cerezo-Llavero, "Calibrating low-cost air quality sensors using multiple arrays of sensors," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2018: IEEE, pp. 1-6.

[16] S. De Vito, G. Di Francia, E. Esposito, S. Ferlito, F. Formisano, and E. Massera, "Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices," *Pattern Recognition Letters,* vol. 136, pp. 264-271, 2020.

[17] S. De Vito *et al.*, "Calibrating chemical multisensory devices for real world applications: An in-depth comparison of quantitative machine learning approaches," *Sensors and Actuators B: Chemical*, vol. 255, pp. 1191-1210, 2018/02/01/ 2018, doi: https://doi.org/10.1016/j.snb.2017.07.155.

[18] W. Jiao *et al.*, "Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States," *Atmospheric Measurement Techniques*, vol. 9, no. 11, pp. 5281-5292, 2016.

[19] J. Hofman, M. Nikolaou, S. P. Shantharam, C. Stroobants, S. Weijs, and V. P. La Manna, "Distant calibration of low-cost PM and NO2 sensors; evidence from multiple sensor testbeds," *Atmospheric Pollution Research,* vol. 13, no. 1, p. 101246, 2022/01/01/ 2022, doi: https://doi.org/10.1016/j.apr.2021.101246.

[20] X. Liu, S. Cheng, H. Liu, S. Hu, D. Zhang, and H. Ning, "A survey on gas sensing technology," *Sensors*, vol. 12, no. 7, pp. 9635-9665, 2012.

[21] B. Maag, Z. Zhou, and L. Thiele, "A Survey on Sensor Calibration in Air Pollution Monitoring Deployments," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4857-4870, 2018, doi: 10.1109/JIOT.2018.2853660.

[22] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 855-873, 2017.

[23] R. S. Sinha, Y. Wei, and S.-H. Hwang, "A survey on LPWA technology: LoRa and NB-IoT," *ICT Express*, vol. 3, no. 1, pp. 14-21, 2017, doi: https://doi.org/10.1016/j.icte.2017.03.004.

[24] W. Yi, K. Lo, T. Mak, K. Leung, Y. Leung, and M. Meng, "A Survey of Wireless Sensor Network Based Air Pollution Monitoring Systems," *Sensors,* vol. 15, no. 12, p. 29859, 2015. [Online]. Available: http://www.mdpi.com/1424-8220/15/12/29859.

[25] M. I. Mead *et al.*, "The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks," *Atmospheric Environment,* vol. 70, pp. 186-203, 2013.

[26] C. Borrego *et al.*, "Assessment of air quality microsensors versus reference methods: The EuNetAir Joint Exercise–Part II," *Atmospheric environment,* vol. 193, pp. 127-142, 2018.

[27] M. Gao, J. Cao, and E. Seto, "A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2. 5 in Xi'an, China," *Environmental pollution,* vol. 199, pp. 56-65, 2015.

[28] O. A. Popoola, G. B. Stewart, M. I. Mead, and R. L. Jones, "Development of a baseline-temperature correction methodology for electrochemical sensors and its implications for long-term stability," *Atmospheric Environment*, vol. 147, pp. 330-343, 2016.

[29] A. E. Fazziki, D. Benslimane, A. Sadiq, J. Ouarzazi, and M. Sadgal, "An Agent Based Traffic Regulation System for the Roadside Air Quality Control," *IEEE Access*, vol. 5, pp. 13192-13201, 2017, doi: 10.1109/ACCESS.2017.2725984.

[30] L. Spinelle, M. Gerboles, M. G. Villani, M. Aleixandre, and F. Bonavitacola, "Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide," *Sensors and Actuators B: Chemical*, vol. 215, pp. 249-257, 2015.

[31] L. Spinelle, M. Gerboles, M. G. Villani, M. Aleixandre, and F. Bonavitacola, "Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. Part B: NO, CO and CO2," *Sensors and Actuators B: Chemical*, vol. 238, pp. 706-715, 2017.

[32] Q. Jiang *et al.*, "Citizen sensing for improved urban environmental monitoring," *Journal of Sensors*, vol. 2016, 2016.

[33] L. Sun *et al.*, "Development and application of a next generation air sensor network for the Hong Kong marathon 2015 air quality monitoring," *Sensors*, vol. 16, no. 2, p. 211, 2016.

[34] C. Borrego *et al.*, "Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise," *Atmospheric Environment*, vol. 147, pp. 246-263, 2016/12/01/ 2016, doi: https://doi.org/10.1016/j.atmosenv.2016.09.050.

[35] S. Brienza, A. Galli, G. Anastasi, and P. Bruschi, "A Low-Cost Sensing System for Cooperative Air Quality Monitoring in Urban Areas," *Sensors*, vol. 15, no. 6, p. 12242, 2015. [Online]. Available: http://www.mdpi.com/1424-8220/15/6/12242.

[36] S. De Vito, M. Piga, L. Martinotto, and G. Di Francia, "CO, NO2 and NOx urban pollution monitoring with on-field calibrated electronic nose by automatic bayesian regularization," *Sensors and Actuators B: Chemical*, vol. 143, no. 1, pp. 182-191, 2009/12/04/ 2009, doi: https://doi.org/10.1016/j.snb.2009.08.041.

[37] X. Fang and I. Bate, "Using multi-parameters for calibration of low-cost sensors in urban environment," *networks*, vol. 7, p. 33, 2017.

[38] S.-C. Hu, Y.-C. Wang, C.-Y. Huang, and Y.-C. Tseng, "Measuring air quality in city areas by vehicular wireless sensor networks," *Journal of Systems and Software*, vol. 84, no. 11, pp. 2005-2012, 2011/11/01/ 2011, doi: https://doi.org/10.1016/j.jss.2011.06.043.

[39] N. Kularatna and B. H. Sudantha, "An Environmental Air Pollution Monitoring System Based on the IEEE 1451 Standard for Low Cost Requirements," *IEEE Sensors Journal*, vol. 8, no. 4, pp. 415-422, 2008, doi: 10.1109/JSEN.2008.917477.

[40] Y. Wang and G. Chen, "Efficient Data Gathering and Estimation for Metropolitan Air Quality Monitoring by Using Vehicular Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7234-7248, 2017, doi: 10.1109/TVT.2017.2655084.

[41] I. Heimann *et al.*, "Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors," *Atmospheric Environment*, vol. 113, pp. 10-19, 2015/07/01/ 2015, doi: https://doi.org/10.1016/j.atmosenv.2015.04.057.

[42] R. Piedrahita *et al.*, "The next generation of low-cost personal air quality sensors for quantitative exposure monitoring," *Atmospheric Measurement Techniques*, vol. 7, no. 10, pp. 3325-3336, 2014.

[43] F. Concas *et al.*, "Low-cost outdoor air quality monitoring and sensor calibration: A survey and critical analysis," *ACM Transactions on Sensor Networks (TOSN)*, vol. 17, no. 2, pp. 1-44, 2021.

[44] F. Karagulian *et al.*, "Review of the performance of low-cost sensors for air quality monitoring," *Atmosphere*, vol. 10, no. 9, p. 506, 2019.

[45] J. E. Thompson, "Crowd-sourced air quality studies: A review of the literature & portable sensors," *Trends in Environmental Analytical Chemistry*, vol. 11, pp. 23-34, 2016/07/01/ 2016, doi: https://doi.org/10.1016/j.teac.2016.06.001.

[46] L. Spinelle, M. Gerboles, G. Kok, S. Persijn, and T. Sauerwald, "Review of Portable and Low-Cost Sensors for the Ambient Air Monitoring of Benzene and Other Volatile Organic Compounds," *Sensors*, vol. 17, no. 7, p. 1520, 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/7/1520.

[47] R. Baron and J. Saffell, "Amperometric gas sensors as a low cost emerging technology platform for air quality monitoring applications: a review," *ACS sensors*, vol. 2, no. 11, pp. 1553-1566, 2017.

[48] S. De Vito, E. Massera, M. Piga, L. Martinotto, and G. Di Francia, "On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario," *Sensors and Actuators B: Chemical*, vol. 129, no. 2, pp. 750-757, 2008.

[49] D. H. Hagan *et al.*, "Calibration and assessment of electrochemical air quality sensors by co-location with regulatory-grade instruments," *Atmospheric Measurement Techniques*, vol. 11, no. 1, pp. 315-328, 2018. [Online]. Available: http://hdl.handle.net/1721.1/114971

[50] E. Esposito, S. De Vito, M. Salvato, V. Bright, R. L. Jones, and O. Popoola, "Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems," *Sensors and Actuators B: Chemical*, vol. 231, pp. 701-713, 2016/08/01/ 2016, doi: https://doi.org/10.1016/j.snb.2016.03.038.

[51] Y. Wang, J. Li, H. Jing, Q. Zhang, J. Jiang, and P. Biswas, "Laboratory evaluation and calibration of three low-cost particle sensors for particulate matter measurement," *Aerosol Science and Technology*, vol. 49, no. 11, pp. 1063-1077, 2015.

[52] G. S. W. Hagler, R. Williams, V. Papapostolou, and A. Polidori, "Air Quality Sensors and Data Adjustment Algorithms: When Is It No Longer a Measurement?," *Environmental Science & Technology*, vol. 52, no. 10, pp. 5530-5531, 2018/05/15 2018, doi: 10.1021/acs.est.8b01826.

[53] D. B. Topalović, M. D. Davidović, M. Jovanović, A. Bartonova, Z. Ristovski, and M. Jovašević-Stojanović, "In search of an optimal in-field calibration method of

low-cost gas sensors for ambient air pollutants: Comparison of linear, multilinear and artificial neural network approaches," *Atmospheric Environment*, vol. 213, pp. 640-658, 2019.

[54] R. Ariyaratne *et al.*, "Understanding the effect of temperature and relative humidity on sensor sensitivities in field environments and improving the calibration models of multiple electrochemical carbon monoxide (CO) sensors in a tropical environment," *Sensors and Actuators B: Chemical*, vol. 390, p. 133935, 2023.

[55] L. Spinelle, M. Gerboles, and M. Aleixandre, "Performance evaluation of amperometric sensors for the monitoring of O3 and NO2 in ambient air at ppb level," *Procedia engineering*, vol. 120, pp. 480-483, 2015.

[56] L. Spinelle, M. Gerboles, M. Aleixandre, and F. Bonavitacola, "Evaluation of metal oxides sensors for the monitoring of O3 in ambient air at Ppb level," *Chemical Engineering Transactions*, vol. 54, pp. 319-324, 2016.

[57] E. Esposito *et al.*, "Stochastic Comparison of Machine Learning Approaches to Calibration of Mobile Air Quality Monitors," Cham, 2018: Springer International Publishing, in Sensors, pp. 294-302.

[58] E. Esposito, S. De Vito, M. Salvato, G. Fattoruso, and G. Di Francia, "Computational Intelligence for Smart Air Quality Monitors Calibration," Cham, 2017: Springer International Publishing, in Computational Science and Its Applications – ICCSA 2017, pp. 443-454.

[59] M. Badura, P. Batog, A. Drzeniecka-Osiadacz, and P. Modzel, "Low- and Medium-Cost Sensors for Tropospheric Ozone Monitoring - Results of an Evaluation Study in Wroclaw, Poland," *Atmosphere*, vol. 13, no. 4, p. 542, 2022. [Online]. Available: https://www.mdpi.com/2073-4433/13/4/542.

[60] M. Rogulski, A. Badyda, A. Gayer, and J. Reis, "Improving the Quality of Measurements Made by Alphasense NO2 Non-Reference Sensors Using the Mathematical Methods," *Sensors*, vol. 22, no. 10, p. 3619, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/10/3619.

[61] C. Zuidema *et al.*, "Deployment, Calibration, and Cross-Validation of Low-Cost Electrochemical Sensors for Carbon Monoxide, Nitrogen Oxides, and Ozone for an Epidemiological Study," *Sensors*, vol. 21, no. 12, p. 4214, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/12/4214.

[62] R. A. González Rivero *et al.*, "Relevance and Reliability of Outdoor SO2 Monitoring in Low-Income Countries Using Low-Cost Sensors," *Atmosphere*, vol. 14, no. 6, p. 912, 2023.

[63] V. Malyan, V. Kumar, and M. Sahu, "Significance of sources and size distribution on calibration of low-cost particle sensors: Evidence from a field sampling campaign," *Journal of Aerosol Science*, vol. 168, p. 106114, 2023/02/01/ 2023, doi: https://doi.org/10.1016/j.jaerosci.2022.106114.

[64] N. Kappelt, H. S. Russell, D. Fessa, K. V. Ryswyk, O. Hertel, and M. S. Johnson, "Particulate air pollution in the Copenhagen metro part 1: Mass concentrations and ventilation," *Environment International*, vol. 171, p. 107621, 2023/01/01/ 2023, doi: https://doi.org/10.1016/j.envint.2022.107621.

[65] J. M. Cordero, R. Borge, and A. Narros, "Using statistical methods to carry out in field calibrations of low cost air quality sensors," *Sensors and Actuators B: Chemical*, vol. 267, pp. 245-254, 2018/08/15/ 2018, doi: https://doi.org/10.1016/j.snb.2018.04.021.

[66] O. Djedidi, M. A. Djeziri, N. Morati, J.-L. Seguin, M. Bendahan, and T. Contaret, "Accurate detection and discrimination of pollutant gases using a temperature modulated MOX sensor combined with feature extraction and support vector classification," *Sensors and Actuators B: Chemical*, vol. 339, p. 129817, 2021/07/15/ 2021, doi: https://doi.org/10.1016/j.snb.2021.129817.

[67] A. Bigi, M. Mueller, S. K. Grange, G. Ghermandi, and C. Hueglin, "Performance of NO, NO2 low cost sensors and three calibration approaches within a real world application," *Atmospheric Measurement Techniques*, vol. 11, no. 6, pp. 3717-3735, 2018.

[68] N. Zimmerman *et al.*, "A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring," *Atmospheric Measurement Techniques*, vol. 11, no. 1, 2018.

[69] E. Bagkis, T. Kassandros, and K. Karatzas, "Learning Calibration Functions on the Fly: Hybrid Batch Online Stacking Ensembles for the Calibration of Low-Cost Air Quality Sensor Networks in the Presence of Concept Drift," *Atmosphere*, vol. 13, no. 3, p. 416, 2022. [Online]. Available: https://www.mdpi.com/2073-4433/13/3/416.

[70] A. S. Bittner, E. S. Cross, D. H. Hagan, C. Malings, E. Lipsky, and A. P. Grieshop, "Performance characterization of low-cost air quality sensors for off-grid deployment in rural Malawi," *Atmospheric Measurement Techniques*, vol. 15, no. 11, pp. 3353-3376, 2022.

[71] C. Malings *et al.*, "Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring," *Atmospheric Measurement Techniques*, vol. 12, no. 2, pp. 903-920, 2019.

[72] A. Wang *et al.*, "Leveraging machine learning algorithms to advance low-cost air sensor calibration in stationary and mobile settings," *Atmospheric Environment*, vol. 301, p. 119692, 2023/05/15/ 2023, doi: https://doi.org/10.1016/j.atmosenv.2023.119692.

[73] J. Á. Martín-Baos, L. Rodriguez-Benitez, R. García-Ródenas, and J. Liu, "IoT based monitoring of air quality and traffic using regression analysis," *Applied Soft Computing*, vol. 115, p. 108282, 2022/01/01/ 2022, doi: https://doi.org/10.1016/j.asoc.2021.108282.

[74] J. P. Sá, H. Chojer, P. T. B. S. Branco, M. C. M. Alvim-Ferraz, F. G. Martins, and S. I. V. Sousa, "Two step calibration method for ozone low-cost sensor: Field experiences with the UrbanSense DCUs," *Journal of Environmental Management*, vol. 328, p. 116910, 2023/02/15/ 2023, doi: https://doi.org/10.1016/j.jenvman.2022.116910.

[75] V. Athira, P. Geetha, R. Vinayakumar, and K. P. Soman, "DeepAirNet: Applying Recurrent Networks for Air Quality Prediction," *Procedia Computer Science*, vol. 132, pp. 1394-1403, 2018/01/01/ 2018, doi: https://doi.org/10.1016/j.procs.2018.05.068.

[76] J. Fonollosa, S. Sheik, R. Huerta, and S. Marco, "Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring," *Sensors and Actuators B: Chemical*, vol. 215, pp. 618-629, 2015.

[77] P. Han *et al.*, "Calibrations of low-cost air pollution monitoring sensors for CO, NO2, O3, and SO2," *Sensors*, vol. 21, no. 1, p. 256, 2021.

[78] L. Xing, "Reliability in Internet of Things: Current status and future perspectives," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6704-6721, 2020.

[79] G. F. Fine, L. M. Cavanagh, A. Afonja, and R. Binions, "Metal Oxide Semi-Conductor Gas Sensors in Environmental Monitoring," *Sensors*, vol. 10, no. 6, pp. 5469-5502, 2010. [Online]. Available: http://www.mdpi.com/1424-8220/10/6/5469.

[80] W. H. Organization, *WHO global air quality guidelines: particulate matter (PM2. 5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization, 2021.

[81] W. H. Organization, *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*. World Health Organization, 2006.

[82] K. Gu, J. Qiao, and W. Lin, "Recurrent Air Quality Predictor Based on Meteorology- and Pollution-Related Factors," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 3946-3955, 2018, doi: 10.1109/TII.2018.2793950.

[83] K.-H. Kim, E. Kabir, and S. Kabir, "A review on the human health impact of airborne particulate matter," *Environment International*, vol. 74, pp. 136-143, 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.envint.2014.10.005.

[84] E. İçöz, F. M. Malik, and K. İçöz, "High spatial resolution IoT based air PM measurement system," *Environmental and Ecological Statistics*, vol. 28, no. 4, pp. 779-792, 2021.

[85] T.-Q. Thach, H. Tsang, P. Cao, and L.-M. Ho, "A novel method to construct an air quality index based on air pollution profiles," *International Journal of Hygiene and Environmental Health*, vol. 221, no. 1, pp. 17-26, 2018/01/01/ 2018, doi: https://doi.org/10.1016/j.ijheh.2017.09.012.

[86] Y. Zhou, F.-J. Chang, L.-C. Chang, I. F. Kao, and Y.-S. Wang, "Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts," *Journal of Cleaner Production*, vol. 209, pp. 134-145, 2019/02/01/ 2019, doi: https://doi.org/10.1016/j.jclepro.2018.10.243.

[87] M. Chen, J. Yang, L. Hu, M. S. Hossain, and G. Muhammad, "Urban Healthcare Big Data System Based on Crowdsourced and Cloud-Based Air Quality Indicators," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 14-20, 2018, doi: 10.1109/MCOM.2018.1700571.

[88] H. Li *et al.*, "Analyzing the impact of heating emissions on air quality index based on principal component regression," *Journal of Cleaner Production*, vol. 171, pp. 1577-1592, 2018/01/10/ 2018, doi: https://doi.org/10.1016/j.jclepro.2017.10.106.

[89] L. Romo-Melo, B. Aristizabal, and M. Orozco-Alzate, "Air-Quality Monitoring in an Urban Area in the Tropical Andes," *IEEE Potentials*, vol. 37, no. 1, pp. 34-39, 2018, doi: 10.1109/MPOT.2017.2714458.

[90] Y. Yang, Z. Zheng, K. Bian, L. Song, and Z. Han, "Real-Time Profiling of Fine-Grained Air Quality Index Distribution Using UAV Sensing," *IEEE Internet of Things Journal,* vol. 5, no. 1, pp. 186-198, 2018, doi: 10.1109/JIOT.2017.2777820.

[91] W. H. Organization. "Ambient (outdoor) air pollution." https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health (accessed 1-Jan-2023, 2023).

[92] K. S. E. Phala, A. Kumar, and G. P. Hancke, "Air Quality Monitoring System Based on ISO/IEC/IEEE 21451 Standards," *IEEE Sensors Journal,* vol. 16, no. 12, pp. 5037-5045, 2016, doi: 10.1109/JSEN.2016.2555935.

[93] M. Jovašević-Stojanović, A. Bartonova, D. Topalović, I. Lazović, B. Pokrić, and Z. Ristovski, "On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter," *Environmental Pollution,* vol. 206, pp. 696-704, 2015.

[94] E. S. Cross *et al.*, "Use of electrochemical sensors for measurement of air pollution: correcting interference response and validating measurements," *Atmos. Meas. Tech.*, vol. 10, no. 9, pp. 3575-3588, 2017, doi: 10.5194/amt-10-3575-2017.

[95] C. Wang, L. Yin, L. Zhang, D. Xiang, and R. Gao, "Metal Oxide Gas Sensors: Sensitivity and Influencing Factors," *Sensors,* vol. 10, no. 3, pp. 2089-2106, 2010. [Online]. Available: http://www.mdpi.com/1424-8220/10/3/2088.

[96] M. Mueller, J. Meyer, and C. Hueglin, "Design of an ozone and nitrogen dioxide sensor unit and its long-term operation within a sensor network in the city of Zurich," *Atmospheric Measurement Techniques,* vol. 10, no. 10, pp. 3783-3799, 2017.

[97] M. Lösch, M. Baumbach, and A. Schütze, "Ozone detection in the ppb-range with improved stability and reduced cross sensitivity," *Sensors and Actuators B: Chemical,* vol. 130, no. 1, pp. 367-373, 2008.

[98] C. R. Martin *et al.*, "Evaluation and environmental correction of ambient CO2 measurements from a low-cost NDIR sensor," *Atmospheric Measurement Techniques,* Article vol. 10, no. 7, pp. 2383-2395, 2017, doi: 10.5194/amt-10-2383-2017.

[99] B. Maag, O. Saukh, D. Hasenfratz, and L. Thiele, "Pre-Deployment Testing, Augmentation and Calibration of Cross-Sensitive Sensors," in *EWSN*, 2016, pp. 169-180.

[100] A. C. Lewis *et al.*, "Evaluating the performance of low cost chemical sensors for air pollution research," *Faraday Discussions,* 10.1039/C5FD00201J vol. 189, no. 0, pp. 85-103, 2016, doi: 10.1039/C5FD00201J.

[101] J. W. Gardner and P. N. Bartlett, "Design of conducting polymer gas sensors: Modelling and experiment," *Synthetic Metals,* vol. 57, no. 1, pp. 3665-3670, 1993/04/12/ 1993, doi: https://doi.org/10.1016/0379-6779(93)90494-H.

[102] H. Bai and G. Shi, "Gas Sensors Based on Conducting Polymers," *Sensors,* vol. 7, no. 3, pp. 267-307, 2007. [Online]. Available: http://www.mdpi.com/1424-8220/7/3/267.

[103] L. Sun, D. Westerdahl, and Z. Ning, "Development and Evaluation of A Novel and Cost-Effective Approach for Low-Cost NO_2 Sensor Drift Correction," *Sensors (Basel)*, vol. 17, no. 8, doi: 10.3390/s17081916.

[104] M. Aleixandre and M. Gerboles, "Review of small commercial sensors for indicative monitoring of ambient gas," *Chemical Engineering Transactions*, vol. 30, 2012.

[105] A. Wisitsoraat, A. Tuantranont, E. Comini, G. Sberveglieri, and W. Wlodarski, "Characterization of n-type and p-type semiconductor gas sensors based on NiOx doped TiO2 thin films," *Thin Solid Films*, vol. 517, no. 8, pp. 2775-2780, 2009/02/27/ 2009, doi: https://doi.org/10.1016/j.tsf.2008.10.090.

[106] N. Yamazoe and K. Shimanoe, "Theory of power laws for semiconductor gas sensors," *Sensors and Actuators B: Chemical*, vol. 128, no. 2, pp. 566-573, 2008/01/15/ 2008, doi: https://doi.org/10.1016/j.snb.2007.07.036.

[107] K. J. Choi and H. W. Jang, "One-Dimensional Oxide Nanostructures as Gas-Sensing Materials: Review and Issues," *Sensors*, vol. 10, no. 4, pp. 4083-4099, 2010. [Online]. Available: http://www.mdpi.com/1424-8220/10/4/4083.

[108] U. Hoefer, H. Böttner, A. Felske, G. Kühner, K. Steiner, and G. Sulz, "Thin-film SnO2 sensor arrays controlled by variation of contact potential—a suitable tool for chemometric gas mixture analysis in the TLV range," *Sensors and Actuators B: Chemical*, vol. 44, no. 1, pp. 429-433, 1997/10/01/ 1997, doi: https://doi.org/10.1016/S0925-4005(97)00230-X.

[109] J. Kukkola *et al.*, "Gas sensors based on anodic tungsten oxide," *Sensors and Actuators B: Chemical*, vol. 153, no. 2, pp. 293-300, 2011/04/20/ 2011, doi: https://doi.org/10.1016/j.snb.2010.10.043.

[110] K. T. Ng, F. Boussaid, and A. Bermak, "A CMOS Single-Chip Gas Recognition Circuit for Metal Oxide Gas Sensor Arrays," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 7, pp. 1569-1580, 2011, doi: 10.1109/TCSI.2011.2143090.

[111] C. H. Kwon *et al.*, "Multi-layered thick-film gas sensor array for selective sensing by catalytic filtering technology," *Sensors and Actuators B: Chemical*, vol. 65, no. 1, pp. 327-330, 2000/06/30/ 2000, doi: https://doi.org/10.1016/S0925-4005(99)00426-8.

[112] W. Zeng, T. Liu, Z. Wang, S. Tsukimoto, M. Saito, and Y. Ikuhara, "Selective Detection of Formaldehyde Gas Using a Cd-Doped TiO2-SnO2 Sensor," *Sensors*, vol. 9, no. 11, pp. 9029-9038, 2009. [Online]. Available: http://www.mdpi.com/1424-8220/9/11/9029.

[113] C. Baratto, E. Comini, G. Faglia, G. Sberveglieri, M. Zha, and A. Zappettini, "Metal oxide nanocrystals for gas sensing," *Sensors and Actuators B: Chemical*, vol. 109, no. 1, pp. 2-6, 2005/08/24/ 2005, doi: https://doi.org/10.1016/j.snb.2005.03.091.

[114] W. Ding, R. Hayashi, J. Suehiro, G. Zhou, K. Imasaka, and M. Hara, "Calibration methods of carbon nanotube gas sensor for partial discharge detection in SF/sub 6," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 13, no. 2, pp. 353-361, 2006, doi: 10.1109/TDEI.2006.1624280.

[115] C. Wongchoosuk, A. Wisitsoraat, D. Phokharatkul, A. Tuantranont, and T. Kerdcharoen, "Multi-walled carbon nanotube-doped tungsten oxide thin films for hydrogen gas sensing," *Sensors*, vol. 10, no. 8, pp. 7705-7715, 2010.

[116] D. Zeng, Y. Hu, J. Liu, Z. Zhao, and S. Gao, "Soft sensing of coal moisture," *Measurement*, vol. 60, pp. 231-239, 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.measurement.2014.09.080.

[117] M. S. McKeown, S. Trabelsi, S. O. Nelson, and E. W. Tollner, "Microwave sensing of moisture in flowing biomass pellets," *Biosystems Engineering*, vol. 155, pp. 152-160, 2017/03/01/ 2017, doi: https://doi.org/10.1016/j.biosystemseng.2016.12.009.

[118] J. Lee *et al.*, "Low power consumption solid electrochemical-type micro CO2 gas sensor," *Sensors and Actuators B: Chemical*, vol. 248, pp. 957-960, 2017/09/01/ 2017, doi: https://doi.org/10.1016/j.snb.2017.02.040.

[119] X. Pang, M. D. Shaw, A. C. Lewis, L. J. Carpenter, and T. Batchellier, "Electrochemical ozone sensors: A miniaturised alternative for ozone measurements in laboratory experiments and air-quality monitoring," *Sensors and Actuators B: Chemical*, vol. 240, pp. 829-837, 2017.

[120] P. Wei *et al.*, "Impact Analysis of Temperature and Humidity Conditions on Electrochemical Sensor Response in Ambient Air Quality Monitoring," *Sensors*, vol. 18, no. 2, p. 59, 2018. [Online]. Available: http://www.mdpi.com/1424-8220/18/2/59.

[121] C. Charlton, B. Temelkuran, G. Dellemann, and B. Mizaikoff, "Midinfrared sensors meet nanotechnology: Trace gas sensing with quantum cascade lasers inside photonic band-gap hollow waveguides," *Applied Physics Letters*, vol. 86, no. 19, p. 194102, 2005, doi: 10.1063/1.1925777.

[122] K. Liu *et al.*, "Investigation of Wavelength Modulation and Wavelength Sweep Techniques in Intracavity Fiber Laser for Gas Detection," *Journal of Lightwave Technology*, vol. 29, no. 1, pp. 15-21, 2011, doi: 10.1109/JLT.2010.2091943.

[123] H. Y. Ryu, W.-K. Lee, H. S. Moon, and H. S. Suh, "Tunable erbium-doped fiber ring laser for applications of infrared absorption spectroscopy," *Optics Communications*, vol. 275, no. 2, pp. 379-384, 2007/07/15/ 2007, doi: https://doi.org/10.1016/j.optcom.2007.03.034.

[124] K. Liu *et al.*, "Wavelength Sweep of Intracavity Fiber Laser for Low Concentration Gas Detection," *IEEE Photonics Technology Letters*, vol. 20, no. 18, pp. 1515-1517, 2008, doi: 10.1109/LPT.2008.928526.

[125] J. Uotila, "Comparison of infrared sources for a differential photoacoustic gas detection system," *Infrared Physics & Technology*, vol. 51, no. 2, pp. 122-130, 2007/10/01/ 2007, doi: https://doi.org/10.1016/j.infrared.2007.05.001.

[126] J. P. Dakin, M. J. Gunning, P. Chambers, and Z. J. Xin, "Detection of gases by correlation spectroscopy," *Sensors and Actuators B: Chemical*, vol. 90, no. 1, pp. 124-131, 2003/04/20/ 2003, doi: https://doi.org/10.1016/S0925-4005(03)00043-1.

[127] A. Cicek *et al.*, "Gas sensing through evanescent coupling of spoof surface acoustic waves," *Sensors and Actuators B: Chemical*, vol. 288, pp. 259-265, 2019/06/01/ 2019, doi: https://doi.org/10.1016/j.snb.2019.02.119.

[128] K.-A. Saalbach, J. Twiefel, and J. Wallaschek, "Self-sensing cavitation detection in ultrasound-induced acoustic cavitation," *Ultrasonics*, vol. 94, pp. 401-410, 2019/04/01/ 2019, doi: https://doi.org/10.1016/j.ultras.2018.06.016.

[129] M. Li *et al.*, "Colloidal quantum dot-based surface acoustic wave sensors for NO2-sensing behavior," *Sensors and Actuators B: Chemical*, vol. 287, pp. 241-249, 2019/05/15/ 2019, doi: https://doi.org/10.1016/j.snb.2019.02.042.

[130] M. Bauersfeld, C. Peter, J. Wollenstein, M. Bucking, J. Bruckert, and J. Steinhanses, "Low-cost gas chromatography with gas sensor array for rapid tests in food industry processes," in *IEEE SENSORS*, 26-29 Oct. 2008 2008, pp. 1167-1170, doi: 10.1109/ICSENS.2008.4716649.

[131] I. G. Kolomnikov, A. M. Efremov, T. I. Tikhomirova, N. M. Sorokina, and Y. A. Zolotov, "Early stages in the history of gas chromatography," *Journal of Chromatography A*, vol. 1537, pp. 109-117, 2018/02/16/ 2018, doi: https://doi.org/10.1016/j.chroma.2018.01.006.

[132] S. I. MacNaughton and S. Sonkusale, "Gas Analysis System on Chip With Integrated Diverse Nanomaterial Sensor Array," *IEEE Sensors Journal*, vol. 15, no. 6, pp. 3500-3506, 2015, doi: 10.1109/JSEN.2015.2391181.

[133] C. Tsamis *et al.*, "Hydrogen catalytic oxidation reaction on Pd-doped porous silicon," *IEEE Sensors Journal*, vol. 2, no. 2, pp. 89-95, 2002, doi: 10.1109/JSEN.2002.1000248.

[134] N.-H. Park, T. Akamatsu, T. Itoh, N. Izu, and W. Shin, "Calorimetric Thermoelectric Gas Sensor for the Detection of Hydrogen, Methane and Mixed Gases," *Sensors*, vol. 14, no. 5, pp. 8350-8362, 2014. [Online]. Available: http://www.mdpi.com/1424-8220/14/5/8350.

[135] P. Kirchner *et al.*, "Thin-film calorimetric H2O2 gas sensor for the validation of germicidal effectivity in aseptic filling processes," *Sensors and Actuators B: Chemical*, vol. 154, no. 2, pp. 257-263, 2011/06/20/ 2011, doi: https://doi.org/10.1016/j.snb.2010.01.058.

[136] J. G. Monroy, A. Lilienthal, J. L. Blanco, J. González-Jimenez, and M. Trincavelli, "Calibration of MOX gas sensors in open sampling systems based on Gaussian Processes," in *SENSORS, 2012 IEEE*, 2012: IEEE, pp. 1-4.

[137] J. G. Monroy, A. J. Lilienthal, J.-L. Blanco, J. Gonzalez-Jimenez, and M. Trincavelli, "Probabilistic gas quantification with MOX sensors in Open Sampling Systems—A Gaussian Process approach," *Sensors and Actuators B: Chemical*, vol. 188, pp. 298-312, 2013.

[138] W. Tsujita, A. Yoshino, H. Ishida, and T. Moriizumi, "Gas sensor network for air-pollution monitoring," *Sensors and Actuators B: Chemical*, vol. 110, no. 2, pp. 304-311, 2005/10/14/ 2005, doi: https://doi.org/10.1016/j.snb.2005.02.008.

[139] P. Peterson *et al.*, "Practical use of metal oxide semiconductor gas sensors for measuring nitrogen dioxide and ozone in urban environments," *Sensors*, vol. 17, no. 7, p. 1653, 2017.

[140] J. Gonzalez-Jimenez, J. G. Monroy, and J. L. Blanco, "The multi-chamber electronic nose—An improved olfaction sensor for mobile robotics," *Sensors*, vol. 11, no. 6, pp. 6145-6164, 2011.

[141] J. R. Stetter and J. Li, "Amperometric gas sensors a review," *Chemical reviews,* vol. 108, no. 2, pp. 352-366, 2008.

[142] T. Kida, A. Nishiyama, M. Yuasa, K. Shimanoe, and N. Yamazoe, "Highly sensitive NO2 sensors using lamellar-structured WO3 particles prepared by an acidification method," *Sensors and Actuators B: Chemical,* vol. 135, no. 2, pp. 568-574, 2009.

[143] R. J. Katulski, J. Namieśnik, J. Stefański, J. Sadowski, W. Wardencki, and K. Szymańska, "Mobile monitoring system for gaseous air pollution," *Metrology and measurement systems,* vol. 16, no. 4, pp. 667-682, 2009.

[144] M. V. Narayana, D. Jalihal, and S. S. Nagendra, "Establishing a sustainable low-cost air quality monitoring setup: A survey of the state-of-the-art," *Sensors,* vol. 22, no. 1, p. 394, 2022.

[145] R. Laref, E. Losson, A. Sava, and M. Siadat, "Empiric Unsupervised Drifts Correction Method of Electrochemical Sensors for in Field Nitrogen Dioxide Monitoring," *Sensors,* vol. 21, no. 11, p. 3581, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/11/3581.

[146] D. L. Auble and T. P. Meyers, "An open path, fast response infrared absorption gas analyzer for H2O and CO2," *Boundary-Layer Meteorology,* journal article vol. 59, no. 3, pp. 243-256, May 01 1992, doi: 10.1007/bf00119815.

[147] H. T. Miyazaki, T. Kasaya, M. Iwanaga, B. Choi, Y. Sugimoto, and K. Sakoda, "Dual-band infrared metasurface thermal emitter for CO2 sensing," *Applied Physics Letters,* vol. 105, no. 12, p. 121107, 2014, doi: 10.1063/1.4896545.

[148] D. M. Holstius, A. Pillarisetti, K. R. Smith, and E. Seto, "Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California," *Atmos. Meas. Tech.*, vol. 7, no. 4, pp. 1121-1131, 2014, doi: 10.5194/amt-7-1121-2014.

[149] P. J. Dacunto *et al.*, "Determining PM 2.5 calibration curves for a low-cost particle monitor: common indoor residential aerosols," *Environmental Science: Processes & Impacts,* vol. 17, no. 11, pp. 1959-1966, 2015.

[150] N. Masson, R. Piedrahita, and M. Hannigan, "Approach for quantification of metal oxide type semiconductor gas sensors used for ambient air quality monitoring," *Sensors and Actuators B: Chemical,* vol. 208, pp. 339-345, 2015.

[151] A. Shrivastava and V. B. Gupta, "Methods for the determination of limit of detection and limit of quantitation of the analytical methods," *Chronicles of young scientists,* vol. 2, no. 1, p. 21, 2011.

[152] E. Austin, I. Novosselov, E. Seto, and M. G. Yost, "Laboratory evaluation of the Shinyei PPD42NS low-cost particulate matter sensor," *PloS one,* vol. 10, no. 9, pp. 1-17, 2015.

[153] F. Kizel *et al.*, "Node-to-node field calibration of wireless distributed air pollution sensor network," *Environmental Pollution,* vol. 233, pp. 900-909, 2018.

[154] J. Kim, A. A. Shusterman, K. J. Lieschke, C. Newman, and R. C. Cohen, "The berkeley atmospheric CO2 observation network: Field calibration and evaluation of low-cost air quality sensors," *Atmospheric Measurement Techniques,* vol. 11, no. 4, pp. 1937-1946, 2018.

[155] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645-1660, 2013/09/01/ 2013, doi: https://doi.org/10.1016/j.future.2013.01.010.

[156] Z. Idrees, Z. Zou, and L. Zheng, "Edge Computing Based IoT Architecture for Low Cost Air Pollution Monitoring Systems: A Comprehensive System Analysis, Design Considerations & Development," *Sensors*, vol. 18, no. 9, p. 3021, 2018. [Online]. Available: http://www.mdpi.com/1424-8220/18/9/3021.

[157] M. C. Carotta *et al.*, "Nanostructured thick-film gas sensors for atmospheric pollutant monitoring: quantitative analysis on field tests," *Sensors and Actuators B: Chemical*, vol. 76, no. 1, pp. 336-342, 2001/06/01/ 2001, doi: https://doi.org/10.1016/S0925-4005(01)00620-7.

[158] P. Althainz, J. Goschnick, S. Ehrmann, and H. Ache, "Multisensor microsystem for contaminants in air," *Sensors and Actuators B: Chemical*, vol. 33, no. 1-3, pp. 72-76, 1996.

[159] F. Röck, N. Barsan, and U. Weimar, "Electronic nose: current status and future trends," *Chemical reviews*, vol. 108, no. 2, pp. 705-725, 2008.

[160] K. P. Singh, S. Gupta, A. Kumar, and S. P. Shukla, "Linear and nonlinear modeling approaches for urban air quality prediction," *Science of The Total Environment*, vol. 426, pp. 244-255, 2012/06/01/ 2012, doi: https://doi.org/10.1016/j.scitotenv.2012.03.076.

[161] R. M. Balabin and E. I. Lomakina, "Support vector machine regression (SVR/LS-SVM)--an alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data," (in eng), *Analyst*, vol. 136, no. 8, pp. 1703-1712, 2011, doi: 10.1039/c0an00387e.

[162] M. Hitchman, N. Cade, T. áKim Gibbs, and N. M. Hedley, "Study of the factors affecting mass transport in electrochemical gas sensors," *Analyst*, vol. 122, no. 11, pp. 1411-1418, 1997.

[163] W. Eugster and G. Kling, "Performance of a low-cost methane sensor for ambient concentration measurements in preliminary studies," *Atmospheric Measurement Techniques*, vol. 5, no. 8, pp. 1925-1934, 2012.

[164] J. H. Sohn, M. Atzeni, L. Zeller, and G. Pioggia, "Characterisation of humidity dependence of a metal oxide semiconductor sensor array using partial least squares," *Sensors and Actuators B: Chemical*, vol. 131, no. 1, pp. 230-235, 2008, doi: https://doi.org/10.1016/j.snb.2007.11.009.

[165] M. Kamionka, P. Breuil, and C. Pijolat, "Calibration of a multivariate gas sensing device for atmospheric pollution measurement," *Sensors and Actuators B: Chemical*, vol. 118, no. 1, pp. 323-327, 2006/10/25/ 2006, doi: https://doi.org/10.1016/j.snb.2006.04.058.

[166] D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele, "Participatory air pollution monitoring using smartphones," *Mobile Sensing*, vol. 1, pp. 1-5, 2012.

[167] S. De Vito, G. Fattoruso, M. Pardo, F. Tortorella, and G. Di Francia, "Semi-supervised learning techniques in artificial olfaction: A novel approach to classification problems and drift counteraction," *IEEE Sensors Journal*, vol. 12, no. 11, pp. 3215-3224, 2012.

[168] P. J. García Nieto, E. F. Combarro, J. J. del Coz Díaz, and E. Montañés, "A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): A case study," *Applied Mathematics and Computation*, vol. 219, no. 17, pp. 8923-8937, 2013/05/01/ 2013, doi: https://doi.org/10.1016/j.amc.2013.03.018.

[169] S. Sheik, S. Marco, R. Huerta, and J. Fonollosa, "Continuous prediction in chemoresistive gas sensors using reservoir computing," *Procedia Engineering*, vol. 87, pp. 843-846, 2014.

[170] O. Saukh, D. Hasenfratz, and L. Thiele, "Reducing multi-hop calibration errors in large-scale mobile sensor networks," in *14th International Conference on Information Processing in Sensor Networks*, 2015: ACM, pp. 274-285.

[171] C. Lin, J. Gillespie, M. Schuder, W. Duberstein, I. Beverland, and M. Heal, "Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide," *Atmospheric Environment*, vol. 100, pp. 111-116, 2015.

[172] M. Gerboles, L. Spinelle, and M. Signorini, "Airsenseur: an open data/software/hardware multi-sensor platform for air quality monitoring. part a: sensor shield," Citeseer, Tech. Rep, 2015.

[173] B. Mijling, Q. Jiang, D. de Jonge, and S. Bocconi, "Practical field calibration of electrochemical NO2 sensors for urban air quality applications," 2017.

[174] Z. Al Barakeh, P. Breuil, N. Redon, C. Pijolat, N. Locoge, and J.-P. Viricelle, "Development of a normalized multi-sensors system for low cost on-line atmospheric pollution detection," *Sensors and Actuators B: Chemical*, vol. 241, pp. 1235-1243, 2017.

[175] B. Maag, Z. Zhou, and L. Thiele, "W-air: Enabling personal air pollution monitoring on wearables," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1-25, 2018.

[176] S. Munir, M. Mayfield, D. Coca, S. A. Jubb, and O. Osammor, "Analysing the performance of low-cost air quality sensors, their drivers, relative benefits and calibration in cities—a case study in Sheffield," *Environmental Monitoring and Assessment*, vol. 191, no. 2, p. 94, 2019/01/22 2019, doi: 10.1007/s10661-019-7231-8.

[177] F. Karagulian *et al.*, "Calibration of AirSensEUR Boxes during a Field Study in the Netherlands," *JRC116324*, 2020.

[178] S. Wang, Y. Hu, J. Burgués, S. Marco, and S.-C. Liu, "Prediction of gas concentration using gated recurrent neural networks," in *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2020: IEEE, pp. 178-182.

[179] I. Vajs, D. Drajic, and Z. Cica, "COVID-19 Lockdown in Belgrade: Impact on Air Pollution and Evaluation of a Neural Network Model for the Correction of Low-Cost Sensors' Measurements," *Applied Sciences*, vol. 11, no. 22, p. 10563, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/22/10563.

[180] D. Wahlborg, M. Björling, and M. Mattsson, "Evaluation of field calibration methods and performance of AQMesh, a low-cost air quality monitor,"

Environmental Monitoring and Assessment, vol. 193, no. 5, p. 251, 2021/04/08 2021, doi: 10.1007/s10661-021-09033-x.

[181] G. Tancev and F. G. Toro, "Variational Bayesian calibration of low-cost gas sensor systems in air quality monitoring," *Measurement: Sensors*, vol. 19, p. 100365, 2022/02/01/ 2022, doi: https://doi.org/10.1016/j.measen.2021.100365.

[182] M. I. Daepp *et al.*, "Eclipse: an end-to-end platform for low-cost, hyperlocal environmental sensing in cities," in *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2022: IEEE, pp. 28-40.

[183] W. A. Jabbar, T. Subramaniam, A. E. Ong, M. I. Shu'Ib, W. Wu, and M. A. de Oliveira, "LoRaWAN-Based IoT System Implementation for Long-Range Outdoor Air Quality Monitoring," *Internet of Things*, vol. 19, p. 100540, 2022/08/01/ 2022, doi: https://doi.org/10.1016/j.iot.2022.100540.

[184] A. Ganji *et al.*, "Design, calibration, and testing of a mobile sensor system for air pollution and built environment data collection: The urban scanner platform," *Environmental Pollution*, vol. 317, p. 120720, 2023/01/15/ 2023, doi: https://doi.org/10.1016/j.envpol.2022.120720.

[185] M. H. Hasan *et al.*, "Unexpected Performance Improvements of Nitrogen Dioxide and Ozone Sensors by Including Carbon Monoxide Sensor Signal," *ACS omega*, vol. 8, no. 6, pp. 5917-5924, 2023.

[186] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural networks for indoor person tracking with infrared sensors," *IEEE Sensors Letters*, vol. 5, no. 1, pp. 1-4, 2021.

[187] H. Cho and S. M. Yoon, "Divide and conquer-based 1D CNN human activity recognition using test data sharpening," *Sensors*, vol. 18, no. 4, p. 1055, 2018.

[188] S. Cavalli and M. Amoretti, "CNN-based multivariate data analysis for bitcoin trend prediction," *Applied Soft Computing*, vol. 101, p. 107065, 2021.

[189] R. R. Kureshi *et al.*, "Data-Driven Techniques for Low-Cost Sensor Selection and Calibration for the Use Case of Air Quality Monitoring," *Sensors*, vol. 22, no. 3, p. 1093, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/1093.

[190] N. E. Johnson, B. Bonczak, and C. E. Kontokosta, "Using a gradient boosting model to improve the performance of low-cost aerosol monitors in a dense, heterogeneous urban environment," *Atmospheric environment*, vol. 184, pp. 9-16, 2018.

[191] C. Persson, P. Bacher, T. Shiga, and H. Madsen, "Multi-site solar power forecasting using gradient boosted regression trees," *Solar Energy*, vol. 150, pp. 423-436, 2017.

[192] N. E. Johnson *et al.*, "Patterns of waste generation: A gradient boosting model for short-term waste prediction in New York City," *Waste management*, vol. 62, pp. 3-11, 2017.

[193] I. F. Ilyas and X. Chu, *Data cleaning*. Morgan & Claypool, 2019.

[194] E. Rahm and H. H. Do, "Data cleaning: Problems and current approaches," *IEEE Data Eng. Bull.*, vol. 23, no. 4, pp. 3-13, 2000.

[195] S. Kappal, "Data normalization using median median absolute deviation MMAD based Z-score for robust predictions vs. min–max normalization," *London Journal of Research in Science: Natural and Formal*, vol. 19, no. 4, p. 10.13140, 2019.

[196] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.", 2022.

[197] P. F. Smith, S. Ganesh, and P. Liu, "A comparison of random forest regression and multiple linear regression for prediction in neuroscience," *Journal of Neuroscience Methods*, vol. 220, no. 1, pp. 85-91, 2013/10/30/ 2013, doi: https://doi.org/10.1016/j.jneumeth.2013.08.024.

[198] T. Dietterich, "Overfitting and undercomputing in machine learning," *ACM computing surveys (CSUR)*, vol. 27, no. 3, pp. 326-327, 1995.

[199] W. Yuchi *et al.*, "Evaluation of random forest regression and multiple linear regression for predicting indoor fine particulate matter concentrations in a highly polluted city," *Environmental Pollution*, vol. 245, pp. 746-753, 2019/02/01/ 2019, doi: https://doi.org/10.1016/j.envpol.2018.11.034.

[200] V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, "Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines," *Ore Geology Reviews*, vol. 71, pp. 804-818, 2015/12/01/ 2015, doi: https://doi.org/10.1016/j.oregeorev.2015.01.001.

[201] J. Cai, K. Xu, Y. Zhu, F. Hu, and L. Li, "Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest," *Applied Energy*, vol. 262, p. 114566, 2020/03/15/ 2020, doi: https://doi.org/10.1016/j.apenergy.2020.114566.

[202] Y. Zhang and A. Haghani, "A gradient boosting method to improve travel time prediction," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 308-324, 2015/09/01/ 2015, doi: https://doi.org/10.1016/j.trc.2015.02.019.

[203] M. Abdi-Khanghah, A. Bemani, Z. Naserzadeh, and Z. Zhang, "Prediction of solubility of N-alkanes in supercritical CO2 using RBF-ANN and MLP-ANN," *Journal of CO2 Utilization*, vol. 25, pp. 108-119, 2018/05/01/ 2018, doi: https://doi.org/10.1016/j.jcou.2018.03.008.

[204] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *International Journal of Artificial Intelligence and Expert Systems*, vol. 1, no. 4, pp. 111-122, 2011.

[205] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020/03/01/ 2020, doi: https://doi.org/10.1016/j.physd.2019.132306.

[206] X. Li *et al.*, "Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation," *Environmental Pollution*, vol. 231, pp. 997-1004, 2017/12/01/ 2017, doi: https://doi.org/10.1016/j.envpol.2017.08.114.

[207] P. Soh, J. Chang, and J. Huang, "Adaptive Deep Learning-Based Air Quality Prediction Model Using the Most Relevant Spatial-Temporal Relations," *IEEE Access*, vol. 6, pp. 38186-38199, 2018, doi: 10.1109/ACCESS.2018.2849820.

[208] P. Wei *et al.*, "Development and evaluation of a robust temperature sensitive algorithm for long term NO2 gas sensor network data correction," *Atmospheric Environment*, vol. 230, p. 117509, 2020.

[209] Q. Wu and H. Lin, "Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and LSTM neural network," *Sustainable Cities and Society*, vol. 50, p. 101657, 2019/10/01/ 2019, doi: https://doi.org/10.1016/j.scs.2019.101657.

[210] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354-377, 2018.

[211] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mechanical systems and signal processing*, vol. 151, p. 107398, 2021.

[212] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067-7075, 2016.

[213] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664-675, 2015.

[214] X. Liu, Q. Zhou, J. Zhao, H. Shen, and X. Xiong, "Fault diagnosis of rotating machinery under noisy environment conditions based on a 1-D convolutional autoencoder and 1-D convolutional neural network," *Sensors*, vol. 19, no. 4, p. 972, 2019.

[215] S. Zhu, X. Lian, H. Liu, J. Hu, Y. Wang, and J. Che, "Daily air quality index forecasting with hybrid models: A case in China," *Environmental Pollution*, vol. 231, pp. 1232-1244, 2017/12/01/ 2017, doi: https://doi.org/10.1016/j.envpol.2017.08.069.

[216] J. Zhu, P. Wu, H. Chen, L. Zhou, and Z. Tao, "A Hybrid Forecasting Approach to Air Quality Time Series Based on Endpoint Condition and Combined Forecasting Model," *International Journal of Environmental Research and Public Health*, vol. 15, no. 9, p. 1941, 2018. [Online]. Available: https://www.mdpi.com/1660-4601/15/9/1941.

[217] P. Wang, Y. Liu, Z. Qin, and G. Zhang, "A novel hybrid forecasting model for PM10 and SO2 daily concentrations," *Science of The Total Environment*, vol. 505, pp. 1202-1212, 2015/02/01/ 2015, doi: https://doi.org/10.1016/j.scitotenv.2014.10.078.

[218] P. Jiang, C. Li, R. Li, and H. Yang, "An innovative hybrid air pollution early-warning system based on pollutants forecasting and Extenics evaluation," *Knowledge-Based Systems*, vol. 164, pp. 174-192, 2019/01/15/ 2019, doi: https://doi.org/10.1016/j.knosys.2018.10.036.

[219] J. Wang, T. Niu, and R. Wang, "Research and Application of an Air Quality Early Warning System Based on a Modified Least Squares Support Vector Machine and a Cloud Model," *International Journal of Environmental Research and Public Health*, vol. 14, no. 3, p. 249, 2017. [Online]. Available: https://www.mdpi.com/1660-4601/14/3/249.

[220] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for machine learning*. Cambridge University Press, 2020.

[221] Alphasense, "Alphasense Application Note: How Electrochemical Gas Sensors Work," *AAN104*, 2018. [Online]. Available: http://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN_104.pdf.

[222] S. Ali, T. Glass, B. Parr, J. Potgieter, and F. Alam, "Low Cost Sensor With IoT LoRaWAN Connectivity and Machine Learning-Based Calibration for Air Pollution Monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-11, 2020.

[223] G. Dong and H. Liu, *Feature engineering for machine learning and data analytics*. CRC press, 2018.

[224] A. Zheng and A. Casari, *Feature engineering for machine learning: principles and techniques for data scientists*. " O'Reilly Media, Inc.", 2018.

CO Dataset 1 results for TTS1 (90:10) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.803	0.554	-0.444	0.029	0.445	0.411
RFR	0.806	0.554	-0.444	0.035	0.446	0.405
GBR	0.805	0.556	-0.445	0.036	0.447	0.405
MLP	0.810	0.551	-0.441	0.041	0.442	0.406
LSTM	0.808	0.545	-0.437	0.026	0.438	0.405
1DCNN	0.811	0.541	-0.435	0.020	0.435	0.397

CO Dataset 1 results for TTS1 (90:10) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.812	0.543	-0.435	0.026	0.436	0.402
RFR	0.837	0.508	-0.407	0.024	0.408	0.372
GBR	0.838	0.506	-0.406	0.023	0.407	0.368
MLP	0.833	0.526	0.422	0.029	0.423	0.388
LSTM	0.838	0.501	-0.403	0.017	0.403	0.375
1DCNN	0.841	0.502	-0.398	0.066	0.404	0.371

CO Dataset 1 results for TTS1 (90:10) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.905	0.384	-0.308	0.021	0.309	0.274
RFR	0.924	0.346	-0.276	0.028	0.278	0.226
GBR	0.923	0.349	-0.279	0.028	0.280	0.233
MLP	0.923	0.350	-0.281	-0.018	0.282	0.232
LSTM	0.925	0.344	-0.276	0.007	0.276	0.227
1DCNN	0.924	0.344	-0.276	0.017	0.276	0.231

CO Dataset 1 results for TTS1 (90:10) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.911	0.373	-0.299	0.020	0.299	0.258
RFR	0.931	0.332	-0.265	0.031	0.266	0.219
GBR	0.941	0.306	-0.245	0.029	0.246	0.213
LSTM	0.929	0.335	-0.267	0.032	0.269	0.229
1DCNN	0.944	0.298	-0.238	0.025	0.240	0.204

CO Dataset 1 results for TTS1 (90:10) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.905	0.386	-0.310	0.018	0.310	0.275
RFR	0.933	0.326	-0.261	0.026	0.262	0.211
GBR	0.929	0.333	-0.267	0.020	0.268	0.212
LSTM	0.932	0.326	-0.261	0.017	0.262	0.210
1DCNN	0.937	0.314	-0.252	0.018	0.253	0.207

CO Dataset 1 results for TTS1 (90:10) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.910	0.373	-0.300	0.017	0.300	0.259
RFR	0.938	0.314	-0.251	0.029	0.252	0.203
GBR	0.945	0.295	-0.236	0.022	0.237	0.195
LSTM	0.934	0.322	-0.257	0.025	0.259	0.214
1DCNN	0.956	0.264	0.211	0.025	0.212	0.180

CO Dataset 1 results for TTS1 (90:10) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.851	0.486	-0.388	0.044	0.390	0.361
RFR	0.883	0.428	-0.342	0.031	0.344	0.309
GBR	0.888	0.420	-0.336	0.036	0.338	0.300
MLP	0.872	0.446	-0.357	0.028	0.358	0.322
LSTM	0.886	0.423	-0.339	0.035	0.340	0.300
1DCNN	0.889	0.419	-0.335	0.037	0.337	0.297

CO Dataset 1 results for TTS1 (90:10) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.901	0.393	-0.315	0.025	0.316	0.279
RFR	0.916	0.363	-0.290	0.032	0.292	0.251
GBR	0.918	0.361	0.289	0.029	0.290	0.242
MLP	0.918	0.364	0.290	0.033	0.292	0.248
LSTM	0.919	0.359	-0.288	0.022	0.289	0.244
1DCNN	0.919	0.354	-0.284	0.008	0.285	0.233

CO Dataset 1 results for TTS1 (90:10) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.915	0.366	-0.293	0.024	0.294	0.256
RFR	0.940	0.308	-0.246	0.028	0.247	0.207
GBR	0.939	0.309	-0.247	0.023	0.248	0.206
MLP	0.937	0.333	0.262	0.055	0.268	0.224
LSTM	0.937	0.314	-0.251	0.019	0.252	0.212
1DCNN	0.938	0.314	-0.250	0.030	0.251	0.208

CO Dataset 1 results for TTS2 (20:80) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.767	0.613	-0.482	0.003	0.482	0.442
RFR	0.770	0.609	-0.479	0.005	0.479	0.437
GBR	0.771	0.609	-0.479	0.004	0.479	0.437
MLP	0.768	0.613	-0.482	0.006	0.482	0.437
LSTM	0.780	0.600	-0.471	0.034	0.472	0.438
1DCNN	0.781	0.598	-0.471	-0.001	0.471	0.430

CO Dataset 1 results for TTS2 (20:80) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.779	0.597	-0.470	0.004	0.470	0.428
RFR	0.781	0.594	-0.468	-0.005	0.468	0.419
GBR	0.781	0.594	-0.467	-0.002	0.467	0.415
MLP	0.786	0.588	-0.463	0.003	0.463	0.417
LSTM	0.798	0.572	-0.450	-0.003	0.450	0.404
1DCNN	0.805	0.564	-0.442	0.043	0.444	0.405

CO Dataset 1 results for TTS2 (20:80) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.882	0.437	-0.344	-0.005	0.344	0.292
RFR	0.899	0.404	-0.318	-0.003	0.318	0.262
GBR	0.897	0.409	-0.322	0.001	0.322	0.266
MLP	0.896	0.415	-0.325	0.033	0.326	0.275
LSTM	0.900	0.405	-0.318	-0.024	0.319	0.257
1DCNN	0.904	0.396	-0.312	-0.014	0.312	0.252

CO Dataset 1 results for TTS2 (20:80) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
RFR	0.903	0.396	-0.311	-0.001	0.311	0.255
LSTM	0.901	0.400	-0.315	0.004	0.315	0.264

CO Dataset 1 results for TTS2 (20:80) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
RFR	0.905	0.392	-0.309	-0.005	0.309	0.250
LSTM	0.904	0.395	-0.311	0.002	0.311	0.249

CO Dataset 1 results for TTS2 (20:80) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.889	0.424	-0.334	-0.004	0.334	0.278
RFR	0.908	0.385	-0.303	-0.003	0.303	0.246
GBR	0.909	0.385	-0.303	-0.002	0.303	0.245
LSTM	0.911	0.380	-0.299	0.000	0.299	0.246
1DCNN	0.918	0.365	-0.287	0.004	0.287	0.236

CO Dataset 1 results for TTS2 (20:80) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.832	0.521	-0.410	0.007	0.410	0.375
RFR	0.851	0.491	-0.387	0.004	0.387	0.340
GBR	0.848	0.496	-0.390	0.012	0.391	0.344
MLP	0.843	0.506	-0.398	0.009	0.398	0.359
LSTM	0.859	0.479	-0.377	0.010	0.377	0.338
1DCNN	0.865	0.467	-0.368	0.013	0.368	0.324

CO Dataset 1 results for TTS2 (20:80) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.876	0.449	-0.353	-0.003	0.353	0.301
RFR	0.891	0.419	-0.330	-0.005	0.330	0.276
GBR	0.887	0.427	-0.336	0.000	0.336	0.282
MLP	0.891	0.429	-0.333	0.054	0.337	0.281
LSTM	0.893	0.418	-0.328	0.018	0.329	0.266
1DCNN	0.898	0.407	-0.320	0.026	0.321	0.262

CO Dataset 1 results for TTS2 (20:80) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.894	0.413	-0.325	-0.002	0.325	0.277
RFR	0.910	0.382	-0.300	-0.002	0.300	0.246
GBR	0.909	0.384	-0.302	-0.003	0.302	0.246
MLP	0.900	0.407	-0.319	0.029	0.320	0.271
LSTM	0.915	0.372	-0.293	-0.004	0.293	0.243
1DCNN	0.920	0.363	-0.286	-0.001	0.286	0.236

NO_2 Dataset 1 results for TTS1 (90:10) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.484	17.786	-0.718	0.027	0.719	14.257
RFR	0.633	14.996	-0.606	0.012	0.606	11.704
GBR	0.628	15.086	-0.609	0.015	0.610	11.703
MLP	0.620	15.476	-0.619	0.087	0.625	12.214
LSTM	0.631	15.187	-0.614	0.006	0.614	11.844
1DCNN	0.633	15.025	-0.606	0.040	0.607	11.737

NO_2 Dataset 1 results for TTS1 (90:10) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.710	13.327	-0.538	0.011	0.539	10.236
RFR	0.820	10.502	-0.424	0.008	0.424	8.056
GBR	0.813	10.709	-0.433	0.005	0.433	8.129
MLP	0.785	11.587	-0.464	0.059	0.468	9.030
LSTM	0.825	10.366	-0.418	0.020	0.419	8.038
1DCNN	0.835	10.059	-0.406	0.007	0.407	7.718

NO_2 Dataset 1 results for TTS1 (90:10) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.785	11.486	-0.464	0.003	0.464	8.930
RFR	0.871	8.886	-0.359	0.006	0.359	6.622
GBR	0.865	9.069	-0.367	0.002	0.367	6.806
MLP	0.855	9.459	-0.382	-0.016	0.382	7.282
LSTM	0.869	8.968	-0.362	0.007	0.362	6.861
1DCNN	0.886	8.445	-0.339	-0.042	0.341	6.343

NO_2 Dataset 1 results for TTS1 (90:10) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.785	11.496	-0.465	0.004	0.465	8.912
RFR	0.909	7.497	15.64	-0.303	-0.005	0.303
GBR	0.911	7.368	-0.298	-0.006	0.298	5.630
LSTM	0.880	8.560	-0.346	0.011	0.346	6.577
1DCNN	0.912	7.355	-0.297	-0.008	0.297	5.626

NO_2 Dataset 1 results for TTS1 (90:10) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.790	11.338	-0.458	0.001	0.458	8.852
RFR	0.884	8.456	-0.342	0.007	0.342	6.346
GBR	0.883	8.472	-0.342	0.012	0.342	6.348
LSTM	0.872	8.836	-0.357	0.011	0.357	6.873
1DCNN	0.885	8.401	0.53	-0.339	0.010	0.339

NO_2 Dataset 1 results for TTS1 (90:10) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.790	11.354	-0.459	0.002	0.459	8.836
RFR	0.915	7.236	-0.292	-0.007	0.292	5.514
GBR	0.923	6.871	-0.278	-0.003	0.278	5.313
LSTM	0.883	8.476	-0.342	0.011	0.343	6.499
1DCNN	0.923	6.888	-0.278	-0.004	0.278	5.342

NO_2 Dataset 1 results for TTS1 (90:10) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.728	12.917	-0.522	0.005	0.522	10.069
RFR	0.835	10.174	-0.411	0.003	0.411	8.080
GBR	0.837	10.115	-0.409	0.009	0.409	7.852
MLP	0.817	10.771	-0.435	0.003	0.435	8.279
LSTM	12.917	10.174	-0.411	0.002	0.411	8.065
1DCNN	0.845	9.867	-0.399	0.009	0.399	7.654

NO_2 Dataset 1 results for TTS1 (90:10) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.757	12.203	-0.493	0.004	0.493	9.279
RFR	0.844	9.796	-0.396	-0.006	0.396	7.237
GBR	0.842	9.832	-0.397	-0.009	0.397	7.284
MLP	0.817	10.598	-0.428	0.005	0.428	8.106
LSTM	0.828	10.260	-0.415	0.011	0.415	7.836
1DCNN	0.843	9.913	-0.399	-0.031	0.401	7.380

NO_2 Dataset 1 results for TTS1 (90:10) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.802	11.007	-0.445	-0.006	0.445	8.558
RFR	0.919	7.100	-0.287	-0.006	0.287	5.381
GBR	0.919	7.024	-0.284	-0.005	0.284	5.371
MLP	0.885	8.499	-0.341	0.038	0.343	6.437
LSTM	0.887	8.362	-0.337	0.027	0.338	6.391
1DCNN	0.918	7.073	-0.286	0.003	0.286	5.400

NO_2 Dataset 1 results for TTS2 (20:80) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.489	17.206	-0.715	-0.042	0.716	13.512
RFR	0.601	15.201	-0.632	-0.038	0.633	11.816
GBR	0.597	15.281	-0.635	-0.038	0.636	11.902
MLP	0.617	14.933	-0.621	-0.023	0.622	11.630
LSTM	0.607	15.155	-0.630	-0.040	0.631	11.881
1DCNN	0.623	14.874	-0.618	-0.037	0.619	11.564

NO_2 Dataset 1 results for TTS2 (20:80) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.677	13.681	-0.569	-0.033	0.569	10.278
RFR	0.763	11.711	-0.487	-0.017	0.487	8.821
GBR	0.760	11.776	-0.490	-0.021	0.490	8.874
MLP	0.758	12.189	-0.493	-0.121	0.507	9.008
LSTM	0.779	11.371	-0.471	-0.048	0.473	8.541
1DCNN	0.790	11.123	-0.462	-0.033	0.463	8.321

NO_2 Dataset 1 results for TTS2 (20:80) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.744	12.187	-0.506	-0.034	0.507	9.122
RFR	0.817	10.291	-0.428	-0.021	0.428	7.613
GBR	0.806	10.593	-0.440	-0.023	0.441	7.899
MLP	0.825	10.082	-0.419	-0.029	0.420	7.359
LSTM	0.811	10.516	-0.437	-0.030	0.438	7.764
1DCNN	0.831	9.894	-0.412	-0.012	0.412	7.192

NO_2 Dataset 1 results for TTS2 (20:80) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
RFR	0.854	9.207	-0.383	-0.020	0.383	6.871
LSTM	0.823	10.125	-0.421	-0.025	0.421	7.546

NO_2 Dataset 1 results for TTS2 (20:80) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
RFR	0.831	9.922	-0.412	-0.024	0.413	7.306
LSTM	0.819	10.240	-0.425	-0.025	0.426	7.639

NO_2 Dataset 1 results for TTS2 (20:80) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.751	12.017	-0.499	-0.038	0.500	8.903
RFR	0.863	8.933	-0.371	-0.022	0.372	6.636
GBR	0.870	8.661	-0.360	-0.016	0.361	6.514
LSTM	0.826	10.043	-0.417	-0.023	0.418	7.514
1DCNN	0.873	8.582	-0.357	-0.020	0.357	6.442

NO_2 Dataset 1 results for TTS2 (20:80) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.699	13.215	-0.549	-0.037	0.550	9.978
RFR	0.788	11.325	-0.471	-0.027	0.471	8.556
GBR	0.786	11.229	-0.467	-0.026	0.467	8.407
MLP	0.760	11.967	-0.497	-0.029	0.498	8.980
LSTM	0.791	11.423	-0.462	0.006	0.462	9.077
1DCNN	0.799	10.902	-0.453	-0.026	0.454	8.140

NO_2 Dataset 1 results for TTS2 (20:80) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.721	12.724	-0.529	-0.034	0.530	9.512
RFR	0.780	11.264	-0.468	-0.020	0.469	8.288
GBR	0.774	11.466	-0.477	-0.009	0.477	8.506
MLP	0.779	11.577	-0.472	-0.096	0.482	8.516
LSTM	0.771	11.574	-0.480	-0.041	0.482	8.548
1DCNN	0.803	10.655	-0.443	0.009	0.444	7.875

NO_2 Dataset 1 results for TTS2 (20:80) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.769	11.595	-0.481	-0.040	0.483	8.565
RFR	0.866	8.854	-0.368	-0.021	0.369	6.582
GBR	0.873	8.586	-0.357	-0.016	0.357	6.432
MLP	0.852	9.420	-0.390	-0.039	0.392	6.944
LSTM	0.836	9.795	-0.406	-0.041	0.408	7.205
1DCNN	0.867	8.807	-0.365	-0.038	0.367	6.468

CO Dataset 2 results for TTS1 (90:10) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.507	0.305	-0.705	-0.019	0.706	0.232
RFR	0.837	0.175	-0.404	0.018	0.405	0.117
GBR	0.842	0.172	-0.397	0.019	0.398	0.116
MLP	0.713	0.252	-0.568	-0.133	0.583	0.183
LSTM	0.837	0.175	-0.405	0.029	0.406	0.120
1DCNN	0.841	0.173	-0.399	0.014	0.399	0.116

CO Dataset 2 results for TTS1 (90:10) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.658	0.254	-0.586	-0.015	0.586	0.196
RFR	0.907	0.132	-0.304	0.017	0.305	0.082
GBR	0.914	0.127	-0.293	0.016	0.293	0.082
MLP	0.755	0.219	-0.501	0.080	0.507	0.168
LSTM	0.909	0.130	-0.301	-0.007	0.301	0.091
1DCNN	0.918	0.124	-0.287	0.016	0.288	0.080

CO Dataset 2 results for TTS1 (90:10) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.710	0.234	-0.540	-0.011	0.540	0.174
RFR	0.910	0.129	-0.299	0.008	0.299	0.076
GBR	0.923	0.120	-0.277	0.009	0.277	0.076
MLP	0.868	0.204	0.459	0.115	0.473	0.139
LSTM	0.924	0.119	-0.275	0.017	0.275	0.083
1DCNN	0.927	0.117	-0.271	0.008	0.271	0.073

CO Dataset 2 results for TTS1 (90:10) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.725	0.227	-0.525	-0.009	0.526	0.168
RFR	0.935	0.110	-0.254	0.011	0.254	0.063
GBR	0.939	0.106	-0.246	0.010	0.246	0.063
LSTM	0.935	0.110	-0.254	0.007	0.255	0.066
1DCNN	0.941	0.105	-0.242	0.011	0.242	0.062

CO Dataset 2 results for TTS1 (90:10) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.710	0.233	-0.539	-0.010	0.539	0.174
RFR	0.917	0.125	-0.288	0.006	0.288	0.073
GBR	0.935	0.110	-0.255	0.010	0.255	0.069
LSTM	0.927	0.117	-0.271	0.007	0.271	0.074
1DCNN	0.934	0.111	-0.257	0.010	0.257	0.068

CO Dataset 2 results for TTS1 (90:10) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.727	0.226	-0.524	-0.008	0.524	0.167
RFR	0.942	0.104	-0.240	0.007	0.240	0.058
GBR	0.953	0.093	-0.216	0.007	0.216	0.054
LSTM	0.937	0.109	-0.251	0.004	0.251	0.070
1DCNN	0.953	0.093	-0.216	0.007	0.216	0.054

CO Dataset 2 results for TTS1 (90:10) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.698	0.238	-0.551	-0.013	0.551	0.179
RFR	0.909	0.131	-0.302	0.008	0.302	0.090
GBR	0.922	0.121	-0.281	0.002	0.281	0.079
MLP	0.831	0.179	-0.414	0.002	0.414	0.117
LSTM	0.823	0.127	-0.422	-0.021	0.423	0.085
1DCNN	0.925	0.119	-0.276	0.001	0.276	0.079

CO Dataset 2 results for TTS1 (90:10) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.604	0.273	-0.631	-0.011	0.631	0.197
RFR	0.879	0.150	-0.347	0.010	0.348	0.095
GBR	0.889	0.144	-0.334	0.015	0.334	0.096
MLP	0.758	0.261	-0.495	-0.346	0.604	0.199
LSTM	0.882	0.150	-0.344	0.038	0.346	0.105
1DCNN	0.893	0.142	-0.328	0.000	0.328	0.098

CO Dataset 2 results for TTS1 (90:10) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.728	0.226	-0.522	-0.004	0.522	0.167
RFR	0.945	0.102	-0.235	0.008	0.235	0.064
GBR	0.947	0.100	-0.231	0.002	0.231	0.062
MLP	0.895	0.154	-0.354	-0.024	0.355	0.109
LSTM	0.943	0.108	0.250	0.002	0.250	0.070
1DCNN	0.952	0.095	-0.219	0.015	0.220	0.061

CO Dataset 2 results for TTS2 (20:80) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.443	0.300	-0.746	-0.025	0.747	0.225
RFR	0.784	0.187	-0.466	-0.009	0.466	0.122
GBR	0.795	0.182	-0.454	-0.011	0.454	0.119
MLP	0.675	0.238	-0.574	0.142	0.592	0.181
LSTM	0.792	0.183	-0.457	0.014	0.457	0.124
1DCNN	0.789	0.185	-0.462	-0.010	0.462	0.121

CO Dataset 2 results for TTS2 (20:80) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.616	0.249	-0.620	-0.012	0.620	0.190
RFR	0.865	0.147	-0.367	0.002	0.367	0.094
GBR	0.869	0.145	-0.362	0.005	0.362	0.095
MLP	0.713	0.217	-0.538	-0.057	0.541	0.160
LSTM	0.870	0.147	-0.365	0.030	0.366	0.106
1DCNN	0.871	0.145	-0.360	0.003	0.360	0.094

CO Dataset 2 results for TTS2 (20:80) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.674	0.229	-0.571	-0.004	0.571	0.170
RFR	0.873	0.143	-0.357	0.008	0.357	0.091
GBR	0.877	0.141	-0.351	0.000	0.351	0.090
MLP	0.842	0.188	0.459	0.090	0.468	0.127
LSTM	0.890	0.134	-0.332	0.032	0.334	0.093
1DCNN	0.887	0.136	-0.336	0.035	0.338	0.094

CO Dataset 2 results for TTS2 (20:80) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
RFR	0.899	0.128	-0.318	0.008	0.319	0.076
LSTM	0.901	0.127	-0.315	0.005	0.315	0.077

CO Dataset 2 results for TTS2 (20:80) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
RFR	0.879	0.140	-0.348	0.009	0.348	0.087
LSTM	0.883	0.138	-0.343	0.002	0.343	0.088

CO Dataset 2 results for TTS2 (20:80) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.690	0.224	-0.557	0.002	0.557	0.163
RFR	0.901	0.126	-0.314	0.007	0.314	0.074
GBR	0.912	0.119	-0.296	0.008	0.296	0.071
LSTM	0.898	0.128	-0.320	0.000	0.320	0.080
1DCNN	0.915	0.117	-0.292	0.006	0.292	0.069

CO Dataset 2 results for TTS2 (20:80) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.653	0.237	-0.589	-0.006	0.589	0.176
RFR	0.867	0.147	-0.366	-0.004	0.366	0.098
GBR	0.774	0.143	-0.477	-0.025	0.478	0.098
MLP	0.797	0.182	-0.454	-0.001	0.454	0.119
LSTM	0.864	0.148	-0.369	-0.002	0.369	0.098
1DCNN	0.876	0.141	-0.352	-0.002	0.352	0.091

CO Dataset 2 results for TTS2 (20:80) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.558	0.267	-0.665	-0.010	0.665	0.194
RFR	0.843	0.159	-0.396	0.001	0.396	0.105
GBR	0.845	0.158	-0.394	0.000	0.394	0.106
MLP	0.681	0.231	0.575	0.042	0.577	0.175
LSTM	0.836	0.164	-0.406	-0.052	0.409	0.115
1DCNN	0.856	0.153	-0.380	-0.020	0.381	0.108

CO Dataset 2 results for TTS2 (20:80) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.694	0.222	-0.553	0.002	0.553	0.163
RFR	0.907	0.122	-0.305	0.008	0.305	0.073
GBR	0.913	0.118	-0.294	0.005	0.294	0.071
MLP	0.862	0.158	-0.377	-0.115	0.394	0.111
LSTM	0.906	0.124	-0.309	-0.002	0.309	0.084
1DCNN	0.918	0.115	-0.286	0.022	0.287	0.075

NO_2 Dataset 2 results for TTS1 (90:10) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.495	8.612	-0.710	-0.029	0.711	6.668
RFR	0.518	8.415	-0.694	-0.034	0.695	6.495
GBR	0.526	8.343	-0.688	-0.029	0.689	6.445
MLP	0.514	8.447	-0.697	-0.009	0.698	6.544
LSTM	0.525	8.357	-0.689	-0.031	0.690	6.455
1DCNN	0.527	8.351	-0.689	-0.023	0.690	6.480

NO_2 Dataset 2 results for TTS1 (90:10) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.559	8.044	-0.664	-0.029	0.664	6.214
RFR	0.658	7.089	-0.585	-0.031	0.585	5.297
GBR	0.678	6.879	-0.568	-0.023	0.568	5.085
MLP	0.617	7.594	-0.627	-0.009	0.627	5.665
LSTM	0.662	7.080	-0.583	-0.039	0.585	5.361
1DCNN	0.661	7.045	-0.582	0.000	0.582	5.318

NO_2 Dataset 2 results for TTS1 (90:10) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.649	7.183	-0.592	-0.038	0.593	5.411
RFR	0.741	6.193	-0.511	-0.027	0.511	4.561
GBR	0.798	5.453	-0.450	-0.020	0.450	3.858
MLP	0.733	6.432	-0.525	-0.077	0.531	4.809
LSTM	0.763	5.896	-0.487	-0.004	0.487	4.275
1DCNN	0.762	6.030	-0.488	-0.097	0.498	4.365

NO_2 Dataset 2 results for TTS1 (90:10) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.649	7.187	-0.592	-0.038	0.593	5.414
RFR	0.764	5.930	-0.489	-0.030	0.490	4.311
GBR	0.826	5.066	-0.418	-0.023	0.418	3.545
LSTM	0.789	5.603	-0.462	-0.027	0.463	4.065
1DCNN	0.821	5.136	-0.423	-0.023	0.424	3.591

NO_2 Dataset 2 results for TTS1 (90:10) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.649	7.184	-0.592	-0.038	0.593	5.420
RFR	0.750	6.088	-0.502	-0.019	0.503	4.489
GBR	0.834	4.938	-0.408	-0.011	0.408	3.542
LSTM	0.778	5.736	-0.473	-0.025	0.474	4.227
1DCNN	0.822	5.114	-0.422	-0.014	0.422	3.660

NO_2 Dataset 2 results for TTS1 (90:10) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.649	7.190	-0.592	-0.039	0.594	5.425
RFR	0.772	5.836	-0.481	-0.023	0.482	4.255
GBR	0.870	4.374	-0.361	-0.011	0.361	3.142
LSTM	0.809	5.342	-0.440	-0.027	0.441	3.904
1DCNN	0.861	4.535	-0.374	-0.016	0.374	3.250

NO_2 Dataset 2 results for TTS1 (90:10) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.572	7.936	-0.655	-0.031	0.655	6.086
RFR	0.697	6.709	-0.553	-0.026	0.554	4.992
GBR	0.701	6.673	-0.551	-0.023	0.551	4.989
MLP	0.688	6.857	-0.566	-0.018	0.566	5.234
LSTM	0.697	6.711	-0.554	-0.026	0.554	4.994
1DCNN	0.707	6.611	-0.545	-0.024	0.546	4.932

NO_2 Dataset 2 results for TTS1 (90:10) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.637	7.314	-0.603	-0.039	0.604	5.514
RFR	0.735	6.259	-0.516	-0.031	0.517	4.604
GBR	0.757	5.991	-0.493	-0.038	0.495	4.328
MLP	0.700	6.720	-0.552	0.059	0.555	5.170
LSTM	0.735	6.269	-0.515	-0.050	0.518	4.539
1DCNN	0.741	6.192	-0.511	-0.027	0.511	4.504

NO_2 Dataset 2 results for TTS1 (90:10) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.666	7.015	-0.578	-0.037	0.579	5.240
RFR	0.784	5.683	-0.469	-0.022	0.469	4.119
GBR	0.868	4.410	-0.364	-0.011	0.364	3.155
MLP	0.778	6.100	-0.478	-0.160	0.504	4.447
LSTM	0.823	5.180	-0.422	-0.071	0.428	3.711
1DCNN	0.857	4.608	-0.380	0.022	0.380	3.330

NO_2 Dataset 2 results for TTS2 (20:80) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.499	8.480	-0.708	0.016	0.708	6.584
RFR	0.506	8.434	-0.704	0.004	0.704	6.503
GBR	0.507	8.427	-0.704	0.007	0.704	6.503
MLP	0.521	8.462	-0.706	-0.013	0.706	6.580
LSTM	0.522	8.301	-0.694	0.008	0.694	6.448
1DCNN	0.518	8.360	-0.694	0.073	0.698	6.569

NO_2 Dataset 2 results for TTS2 (20:80) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.550	8.037	-0.671	0.017	0.671	6.287
RFR	0.618	7.419	-0.619	0.006	0.619	5.547
GBR	0.618	7.412	-0.619	0.003	0.619	5.617
MLP	0.612	7.523	-0.623	0.078	0.628	5.780
LSTM	0.624	7.366	-0.614	0.035	0.615	5.609
1DCNN	0.625	7.338	-0.613	-0.014	0.613	5.537

NO_2 Dataset 2 results for TTS2 (20:80) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.633	7.256	-0.606	0.011	0.606	5.515
RFR	0.723	6.303	-0.526	0.008	0.526	4.627
GBR	0.735	6.184	-0.516	0.011	0.516	4.488
MLP	0.710	6.489	-0.542	-0.017	0.542	4.879
LSTM	0.724	6.340	-0.526	0.061	0.529	4.721
1DCNN	0.730	6.242	-0.521	-0.017	0.521	4.589

NO_2 Dataset 2 results for TTS2 (20:80) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
RFR	0.729	6.234	-0.520	0.004	0.520	4.597
LSTM	0.739	6.125	-0.511	0.009	0.511	4.487

NO_2 Dataset 2 results for TTS2 (20:80) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
RFR	0.729	6.238	-0.521	0.004	0.521	4.652
LSTM	0.723	6.309	-0.527	0.013	0.527	4.733

NO_2 Dataset 2 results for TTS2 (20:80) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.633	7.259	-0.606	0.010	0.606	5.527
RFR	0.745	6.055	-0.506	0.002	0.506	4.506
GBR	0.801	5.349	-0.446	0.013	0.447	3.878
LSTM	0.739	6.115	-0.510	0.011	0.511	4.549
1DCNN	0.795	5.424	-0.453	0.011	0.453	3.942

NO_2 Dataset 2 results for TTS2 (20:80) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.565	7.906	-0.660	0.015	0.660	6.158
RFR	0.651	7.084	-0.591	0.009	0.591	5.405
GBR	0.660	6.991	-0.583	0.019	0.584	5.357
MLP	0.615	7.439	-0.621	0.013	0.621	5.738
LSTM	0.660	7.018	-0.586	0.020	0.586	5.405
1DCNN	0.664	6.953	-0.580	0.017	0.580	5.308

NO_2 Dataset 2 results for TTS2 (20:80) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.623	7.361	-0.614	0.010	0.615	5.582
RFR	0.705	6.511	-0.544	0.001	0.544	4.803
GBR	0.705	6.510	-0.543	0.008	0.543	4.806
MLP	0.675	6.845	-0.571	-0.006	0.571	5.114
LSTM	0.688	6.690	-0.558	-0.019	0.559	4.964
1DCNN	0.699	6.582	-0.549	-0.026	0.549	4.864

NO_2 Dataset 2 results for TTS2 (20:80) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.648	7.110	-0.594	0.009	0.594	5.383
RFR	0.763	5.837	-0.487	0.006	0.487	4.281
GBR	0.803	5.323	-0.444	0.012	0.444	3.862
MLP	0.737	6.156	-0.514	0.016	0.514	4.618
LSTM	0.742	6.103	-0.509	0.022	0.509	4.430
1DCNN	0.783	5.658	-0.466	-0.075	0.472	4.114

CO Dataset 3 results for TTS1 (90:10) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.920	0.075	-0.283	0.039	0.285	0.056
RFR	0.919	0.075	-0.284	0.041	0.287	0.058
GBR	0.926	0.072	-0.274	-0.025	0.275	0.055
MLP	0.926	0.073	-0.273	0.039	0.276	0.056
LSTM	0.922	0.074	-0.280	0.040	0.283	0.057
1DCNN	0.926	0.072	-0.272	0.036	0.274	0.056

CO Dataset 3 results for TTS1 (90:10) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.948	0.060	-0.228	0.023	0.229	0.048
RFR	0.970	0.046	-0.173	0.023	0.175	0.033
GBR	0.972	0.044	-0.167	0.017	0.168	0.033
MLP	0.965	0.049	-0.186	-0.018	0.187	0.039
LSTM	0.971	0.045	-0.171	0.023	0.172	0.033
1DCNN	0.973	0.044	-0.165	0.015	0.166	0.032

CO Dataset 3 results for TTS1 (90:10) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.958	0.054	-0.206	0.009	0.206	0.042
RFR	0.974	0.043	-0.163	0.012	0.163	0.029
GBR	0.980	0.038	0.143	0.005	0.143	0.026
MLP	0.971	0.049	0.186	0.001	0.186	0.037
LSTM	0.978	0.039	-0.147	0.003	0.147	0.026
1DCNN	0.979	0.038	-0.146	0.001	0.146	0.026

CO Dataset 3 results for TTS1 (90:10) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.958	0.054	-0.205	0.009	0.205	0.042
RFR	0.983	0.034	-0.131	0.005	0.131	0.022
GBR	0.991	0.025	0.095	0.003	0.095	0.016
LSTM	0.988	0.029	0.110	0.002	0.111	0.017
1DCNN	0.990	0.026	0.098	-0.003	0.098	0.017

CO Dataset 3 results for TTS1 (90:10) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.958	0.054	-0.206	0.009	0.206	0.042
RFR	0.974	0.042	-0.161	0.013	0.161	0.028
GBR	0.980	0.038	-0.143	0.002	0.143	0.026
LSTM	0.980	0.038	-0.143	0.002	0.143	0.027
1DCNN	0.979	0.038	-0.143	0.002	0.143	0.026

CO Dataset 3 results for TTS1 (90:10) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.958	0.054	-0.205	0.009	0.205	0.042
RFR	0.983	0.034	-0.130	0.007	0.130	0.022
GBR	0.991	0.025	0.096	0.000	0.096	0.016
LSTM	0.990	0.027	0.102	-0.002	0.102	0.019
1DCNN	0.991	0.026	0.098	0.002	0.098	0.017

CO Dataset 3 results for TTS1 (90:10) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.958	0.054	-0.205	0.019	0.206	0.042
RFR	0.976	0.041	-0.155	0.021	0.157	0.030
GBR	0.979	0.039	-0.148	-0.001	0.148	0.019
MLP	0.967	0.049	-0.186	-0.001	0.186	0.022
LSTM	0.978	0.040	-0.152	-0.002	0.152	0.019
1DCNN	0.981	0.037	-0.139	0.000	0.139	0.018

CO Dataset 3 results for TTS1 (90:10) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.939	0.065	-0.246	0.017	0.247	0.050
RFR	0.947	0.061	-0.231	0.019	0.231	0.045
GBR	0.961	0.052	-0.198	0.003	0.198	0.039
MLP	0.951	0.059	-0.223	0.019	0.223	0.045
LSTM	0.975	0.047	0.159	0.084	0.179	0.036
1DCNN	0.974	0.043	0.165	0.006	0.165	0.033

CO Dataset 3 results for TTS1 (90:10) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.963	0.051	-0.193	0.011	0.193	0.040
RFR	0.987	0.030	0.114	0.012	0.114	0.019
GBR	0.992	0.024	0.092	0.001	0.092	0.015
MLP	0.983	0.041	-0.133	0.085	0.158	0.034
LSTM	0.992	0.026	-0.087	-0.050	0.100	0.020
1DCNN	0.993	0.023	0.085	-0.022	0.088	0.017

CO Dataset 3 results for TTS2 (20:80) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.895	0.080	-0.324	0.005	0.324	0.060
RFR	0.890	0.082	-0.332	0.007	0.332	0.061
GBR	0.890	0.082	-0.332	0.005	0.332	0.061
MLP	0.897	0.082	0.325	0.070	0.332	0.064
LSTM	0.894	0.080	-0.326	0.009	0.326	0.060
1DCNN	0.901	0.079	-0.317	0.041	0.320	0.062

CO Dataset 3 results for TTS2 (20:80) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.927	0.067	-0.271	0.005	0.271	0.051
RFR	0.933	0.064	-0.259	0.011	0.259	0.045
GBR	0.936	0.063	-0.253	0.010	0.254	0.046
MLP	0.936	0.063	-0.254	0.012	0.254	0.047
LSTM	0.937	0.062	-0.251	0.009	0.251	0.045
1DCNN	0.937	0.062	-0.250	0.010	0.251	0.045

CO Dataset 3 results for TTS2 (20:80) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.941	0.060	-0.243	-0.001	0.243	0.046
RFR	0.941	0.060	-0.243	0.004	0.243	0.042
GBR	0.954	0.053	-0.214	0.004	0.214	0.039
MLP	0.957	0.056	0.221	-0.054	0.227	0.043
LSTM	0.954	0.053	-0.214	0.003	0.214	0.038
1DCNN	0.967	0.049	0.194	-0.035	0.197	0.036

CO Dataset 3 results for TTS2 (20:80) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
RFR	0.952	0.054	-0.219	0.006	0.219	0.037
LSTM	0.967	0.045	-0.183	0.007	0.183	0.028

CO Dataset 3 results for TTS2 (20:80) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
RFR	0.942	0.060	-0.242	0.008	0.242	0.042
LSTM	0.952	0.054	-0.220	0.002	0.220	0.039

CO Dataset 3 results for TTS2 (20:80) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.941	0.060	-0.243	-0.001	0.243	0.046
RFR	0.951	0.055	-0.221	0.008	0.221	0.038
GBR	0.959	0.050	-0.203	0.000	0.203	0.029
LSTM	0.970	0.043	-0.173	0.007	0.173	0.028
1DCNN	0.969	0.043	-0.175	0.006	0.175	0.028

CO Dataset 3 results for TTS2 (20:80) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.941	0.060	-0.243	0.004	0.243	0.045
RFR	0.957	0.050	-0.206	0.015	0.207	0.034
GBR	0.957	0.051	-0.208	0.010	0.208	0.027
MLP	0.956	0.052	-0.210	0.011	0.210	0.028
LSTM	0.961	0.049	-0.198	0.009	0.198	0.027
1DCNN	0.965	0.047	-0.188	0.014	0.189	0.027

CO Dataset 3 results for TTS2 (20:80) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.920	0.070	-0.284	-0.005	0.284	0.054
RFR	0.929	0.066	-0.266	0.003	0.266	0.047
GBR	0.925	0.068	-0.274	0.005	0.274	0.049
MLP	0.929	0.068	-0.272	0.050	0.277	0.055
LSTM	0.940	0.065	-0.253	0.073	0.263	0.051
1DCNN	0.942	0.063	0.250	0.043	0.254	0.049

CO Dataset 3 results for TTS2 (20:80) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppm)	nCRMSE	nMBE	nRMSE	MAE (ppm)
LR/MLR	0.952	0.054	-0.220	0.000	0.220	0.042
RFR	0.958	0.051	-0.206	0.013	0.206	0.033
GBR	0.964	0.047	-0.189	0.005	0.189	0.026
MLP	0.970	0.048	0.176	-0.081	0.194	0.039
LSTM	0.979	0.036	0.146	-0.014	0.146	0.026
1DCNN	0.986	0.031	0.118	-0.043	0.126	0.023

NO_2 Dataset 3 results for TTS1 (90:10) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.734	7.015	-0.515	-0.014	0.515	5.136
RFR	0.762	6.634	-0.487	-0.017	0.487	4.693
GBR	0.761	6.644	-0.488	-0.022	0.488	4.678
MLP	0.750	6.898	-0.500	-0.082	0.507	4.908
LSTM	0.765	6.635	-0.488	-0.003	0.488	4.719
1DCNN	0.761	6.768	-0.493	-0.068	0.497	4.759

NO_2 Dataset 3 results for TTS1 (90:10) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.792	6.206	-0.455	-0.025	0.456	4.484
RFR	0.873	4.849	-0.356	-0.021	0.356	3.266
GBR	0.877	4.776	-0.351	-0.018	0.351	3.169
MLP	0.853	5.517	-0.388	-0.118	0.405	3.977
LSTM	0.886	4.711	0.342	-0.052	0.346	3.345
1DCNN	0.891	4.567	-0.336	-0.007	0.336	3.039

NO_2 Dataset 3 results for TTS1 (90:10) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.806	5.995	-0.440	-0.022	0.441	4.293
RFR	0.889	4.549	-0.334	-0.020	0.334	3.132
GBR	0.910	4.079	-0.299	-0.017	0.300	2.714
MLP	0.873	4.916	-0.359	-0.038	0.361	3.602
LSTM	0.943	3.317	0.240	0.040	0.244	2.312
1DCNN	0.947	3.219	-0.230	-0.057	0.237	2.278

NO_2 Dataset 3 results for TTS1 (90:10) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.834	5.546	-0.407	-0.014	0.408	3.961
RFR	0.901	4.305	-0.316	-0.017	0.316	2.920
GBR	0.938	3.383	-0.248	-0.015	0.249	2.232
LSTM	0.952	3.084	-0.221	0.051	0.227	2.315
1DCNN	0.947	3.116	-0.229	-0.001	0.229	2.022

NO_2 Dataset 3 results for TTS1 (90:10) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.806	5.989	-0.440	-0.021	0.440	4.269
RFR	0.893	4.474	-0.328	-0.020	0.329	3.060
GBR	0.918	3.896	-0.286	-0.011	0.286	2.620
LSTM	0.947	3.207	-0.232	0.042	0.236	2.343
1DCNN	0.945	3.200	-0.235	0.001	0.235	2.092

NO_2 Dataset 3 results for TTS1 (90:10) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.834	5.537	-0.407	-0.013	0.407	3.915
RFR	0.903	4.277	-0.314	-0.018	0.314	2.888
GBR	0.937	3.408	-0.250	-0.011	0.250	2.202
LSTM	0.960	2.958	-0.202	0.081	0.217	2.282
1DCNN	0.947	3.141	-0.231	0.000	0.231	2.039

NO_2 Dataset 3 results for TTS1 (90:10) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.814	5.876	-0.431	-0.018	0.432	4.217
RFR	0.899	4.360	-0.320	-0.019	0.320	2.962
GBR	0.915	3.986	-0.292	-0.017	0.293	2.755
MLP	0.869	4.960	-0.364	-0.025	0.364	3.479
LSTM	0.925	3.726	-0.273	-0.016	0.274	2.514
1DCNN	0.937	3.418	-0.251	-0.014	0.251	2.304

NO_2 Dataset 3 results for TTS1 (90:10) SC-G (raw LCS + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.768	6.546	-0.481	-0.010	0.481	4.626
RFR	0.835	5.523	-0.406	-0.004	0.406	3.774
GBR	0.836	5.504	-0.404	-0.018	0.404	3.698
MLP	0.823	5.980	-0.426	-0.106	0.439	4.220
LSTM	0.861	5.364	0.385	0.086	0.394	3.744
1DCNN	0.862	5.059	-0.372	-0.006	0.372	3.429

NO_2 Dataset 3 results for TTS1 (90:10) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.834	5.559	-0.408	-0.018	0.409	3.928
RFR	0.906	4.198	-0.308	-0.015	0.308	2.840
GBR	0.933	3.531	-0.259	-0.014	0.259	2.372
MLP	0.921	4.459	0.312	-0.099	0.328	3.413
LSTM	0.958	2.796	-0.204	-0.022	0.205	2.030
1DCNN	0.966	2.521	-0.184	0.020	0.185	1.854

NO_2 Dataset 3 results for TTS2 (20:80) SC1 (raw LCS)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.779	6.433	-0.470	-0.029	0.471	4.787
RFR	0.787	6.323	-0.462	-0.028	0.463	4.704
GBR	0.778	6.453	-0.471	-0.027	0.472	4.825
MLP	0.756	7.493	-0.502	-0.221	0.548	5.598
LSTM	0.797	6.181	-0.452	-0.018	0.452	4.542
1DCNN	0.803	6.128	-0.444	-0.065	0.448	4.484

NO_2 Dataset 3 results for TTS2 (20:80) SC2 (raw LCS + Temp + Hum)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.822	5.766	-0.421	-0.025	0.422	4.269
RFR	0.867	4.982	-0.364	-0.020	0.365	3.557
GBR	0.857	5.168	-0.378	-0.014	0.378	3.713
MLP	0.857	5.399	-0.378	-0.114	0.395	3.947
LSTM	0.874	5.073	-0.369	-0.039	0.371	3.653
1DCNN	0.888	4.585	-0.336	0.004	0.336	3.297

NO_2 Dataset 3 results for TTS2 (20:80) SC3 (or S1) (raw LCS + Temp + Hum + other gases)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.828	5.685	-0.415	-0.028	0.416	4.174
RFR	0.881	4.737	-0.346	-0.023	0.347	3.357
GBR	0.890	4.541	-0.332	-0.012	0.332	3.191
MLP	0.852	5.582	-0.385	0.136	0.409	4.187
LSTM	0.914	4.134	-0.294	0.072	0.303	2.889
1DCNN	0.919	3.930	-0.284	-0.044	0.288	2.816

NO_2 Dataset 3 results for TTS2 (20:80) S2 (raw LCS + Temp + Hum + other gases + N_{days})

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
RFR	0.888	4.606	-0.336	-0.025	0.337	3.241
LSTM	0.913	4.033	-0.295	-0.015	0.295	2.819

NO_2 Dataset 3 results for TTS2 (20:80) S3 (raw LCS + Temp + Hum + other gases + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
RFR	0.882	4.716	-0.344	-0.022	0.345	3.323
LSTM	0.909	4.130	-0.302	-0.014	0.302	2.937

NO_2 Dataset 3 results for TTS2 (20:80) S4 (raw LCS + Temp + Hum + other gases + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.850	5.299	-0.387	-0.019	0.388	3.879
RFR	0.888	4.607	-0.336	-0.023	0.337	3.228
GBR	0.909	4.130	-0.302	-0.014	0.302	2.937
LSTM	0.913	4.036	-0.295	-0.013	0.295	2.834
1DCNN	0.935	3.594	-0.255	0.064	0.263	2.600

NO_2 Dataset 3 results for TTS2 (20:80) SC2T (raw LCS + Temp + Hum + N_{days} + Hour)

Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.846	5.356	-0.392	-0.011	0.392	3.940
RFR	0.888	4.607	-0.337	-0.020	0.337	3.285
GBR	0.898	4.380	-0.320	-0.012	0.321	3.074
MLP	0.877	4.802	-0.351	0.001	0.351	3.322
LSTM	0.901	4.298	-0.315	-0.003	0.315	3.020
1DCNN	0.912	4.055	-0.297	-0.005	0.297	2.868

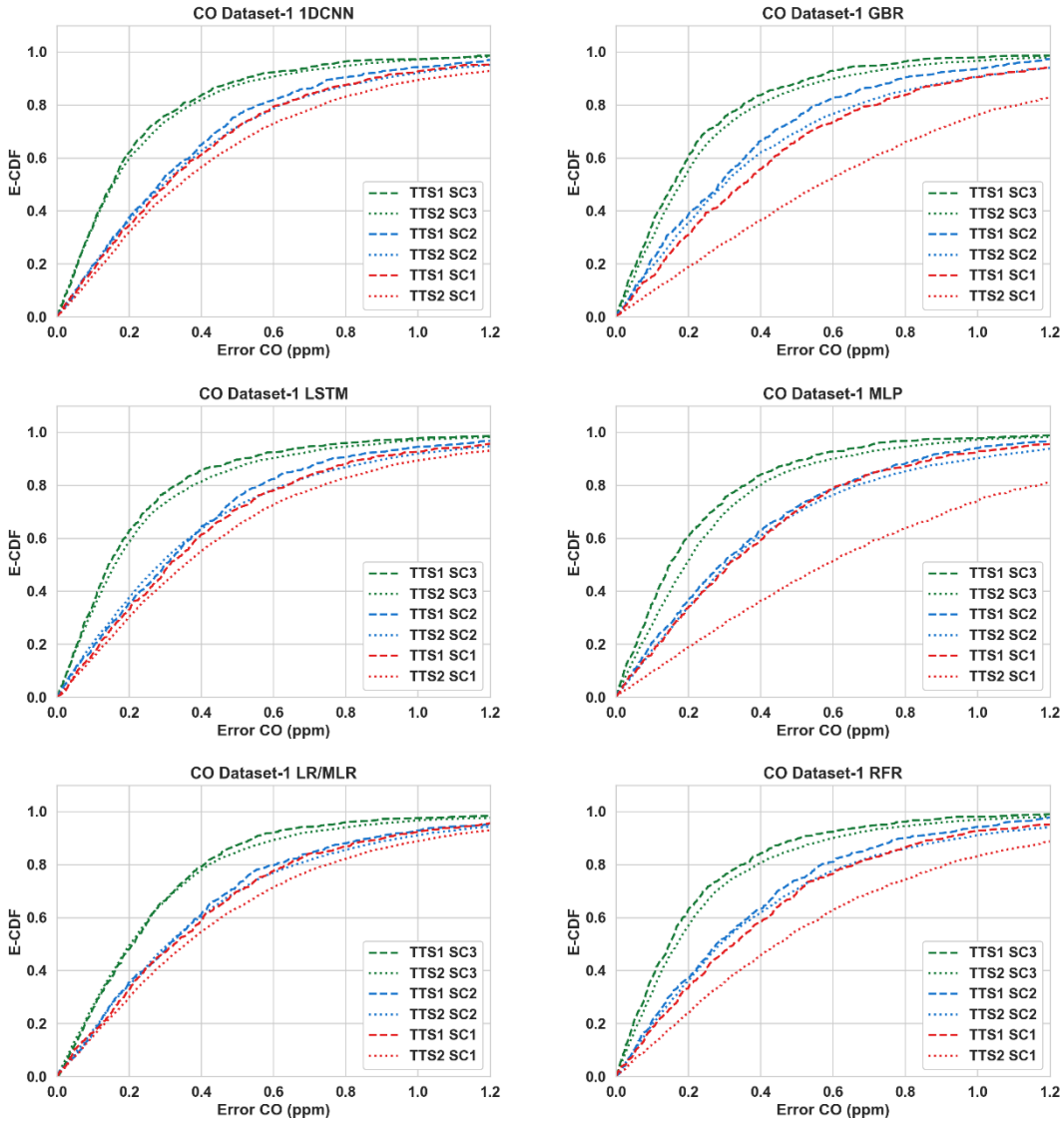
NO_2 Dataset 3 results for TTS2 (20:80) SC-G (raw LCS + other gases)

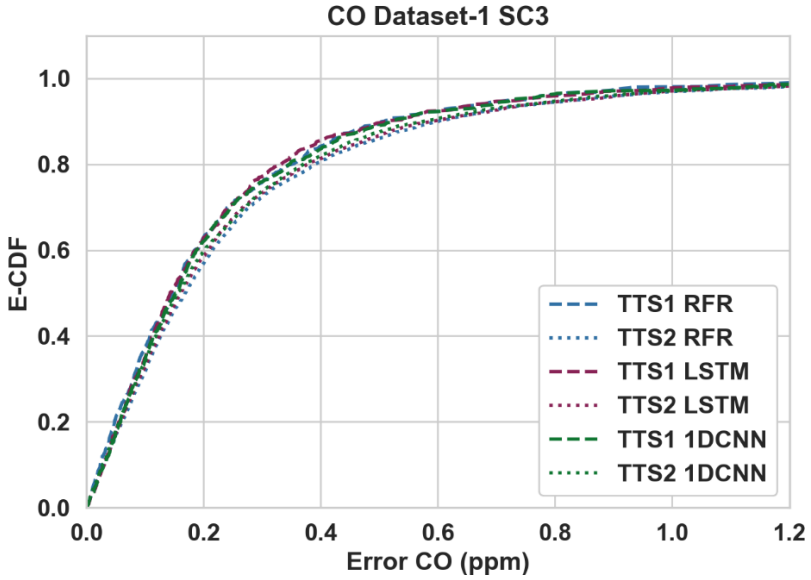
Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.795	6.205	-0.453	-0.036	0.454	4.557
RFR	0.825	5.740	-0.419	-0.033	0.420	4.108
GBR	0.816	5.872	-0.429	-0.021	0.430	4.252
MLP	0.816	6.121	-0.429	-0.129	0.448	4.517
LSTM	0.844	5.469	-0.396	-0.057	0.400	3.841
1DCNN	0.859	5.132	-0.375	-0.023	0.376	3.678

NO_2 Dataset 3 results for TTS2 (20:80) SC-T (SC3 + Year + Month + Date + Week-Day/End + N_{days} + Hour)

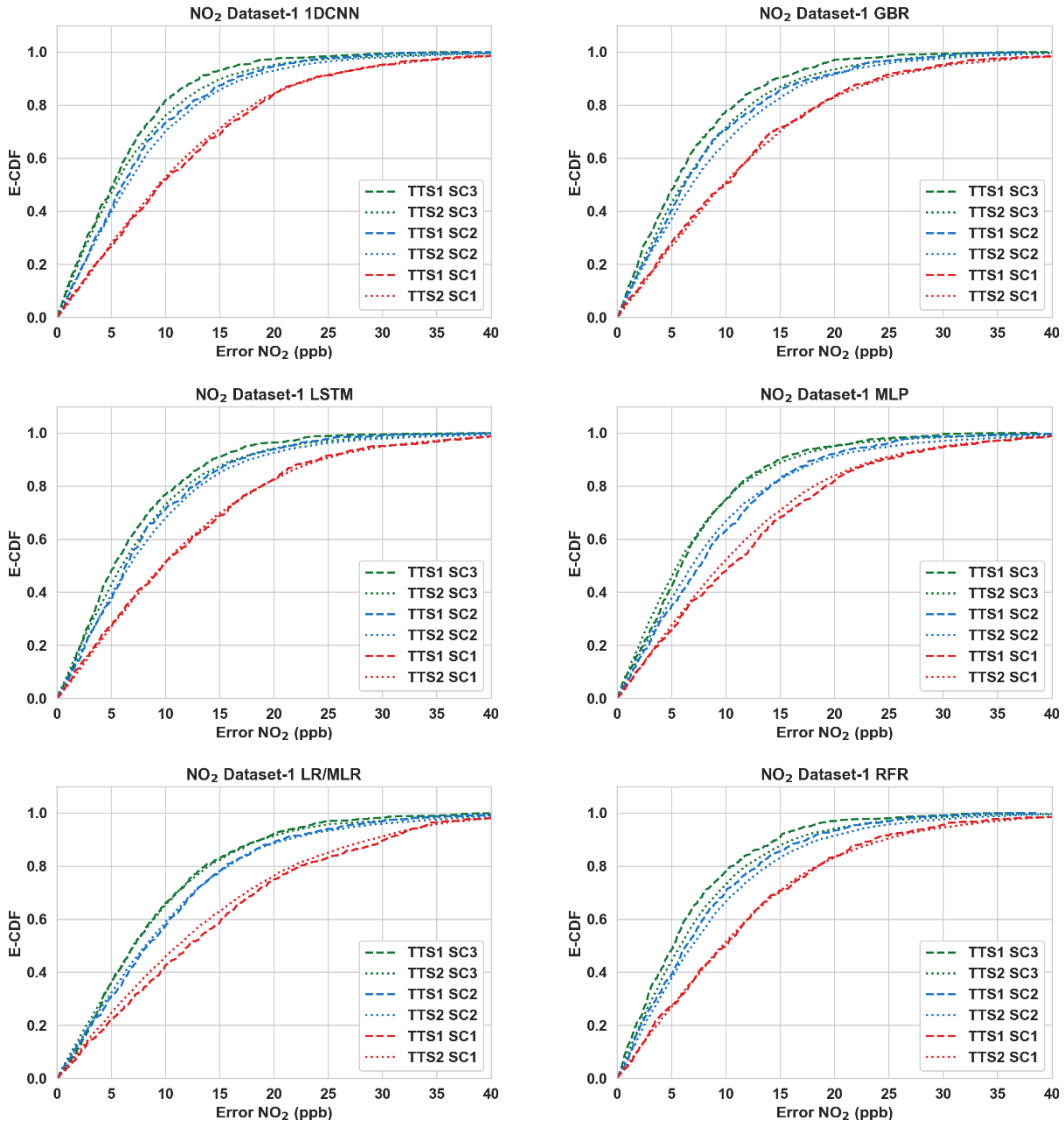
Algorithm	Performance Metrics					
	R^2	RMSE (ppb)	nCRMSE	nMBE	nRMSE	MAE (ppb)
LR/MLR	0.856	5.183	-0.379	-0.017	0.379	3.821
RFR	0.892	4.530	-0.331	-0.022	0.332	3.187
GBR	0.908	4.150	-0.304	-0.011	0.304	2.931
MLP	0.893	4.784	-0.329	-0.120	0.350	3.499
LSTM	0.928	3.754	0.273	-0.032	0.275	2.636
1DCNN	0.935	3.594	-0.255	0.064	0.263	2.600

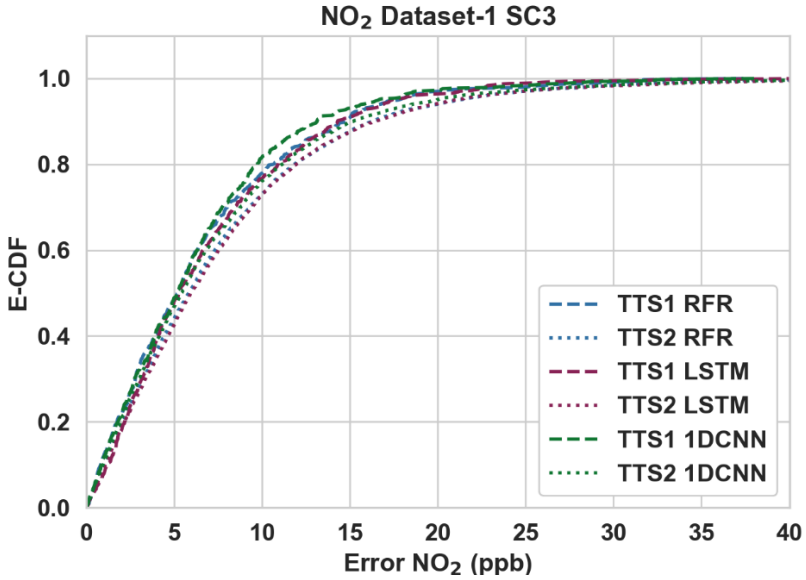
E-CDFs of CO Dataset 1 for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



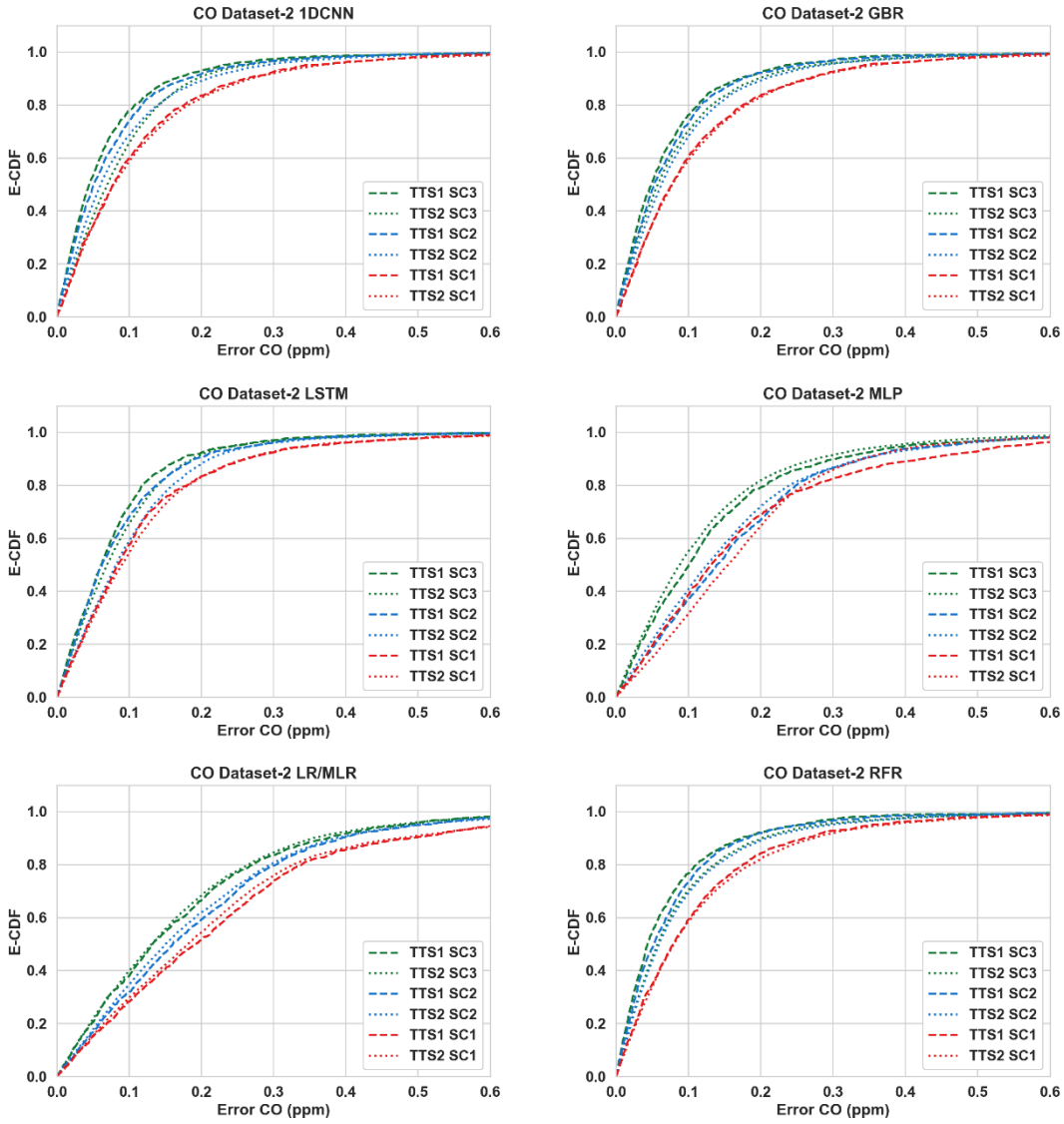


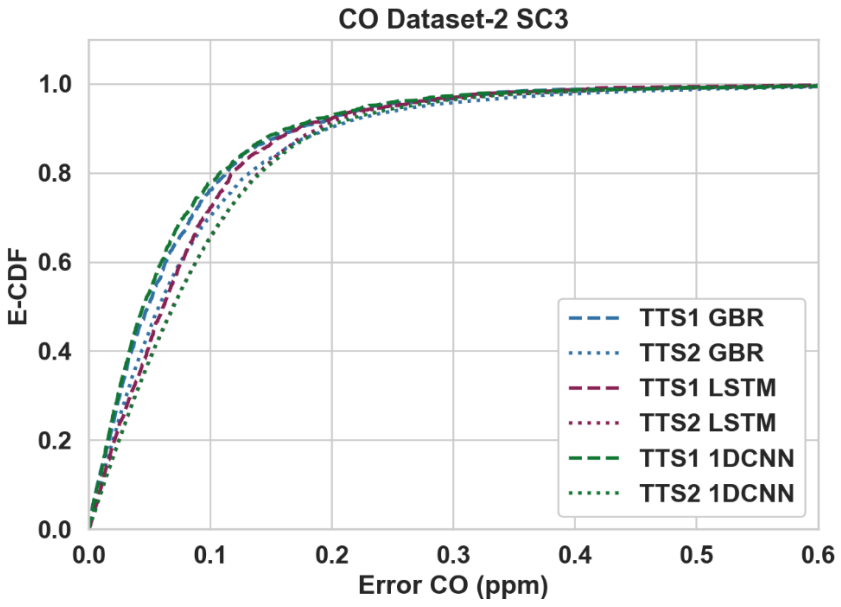
E-CDFs of NO_2 Dataset 1 for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



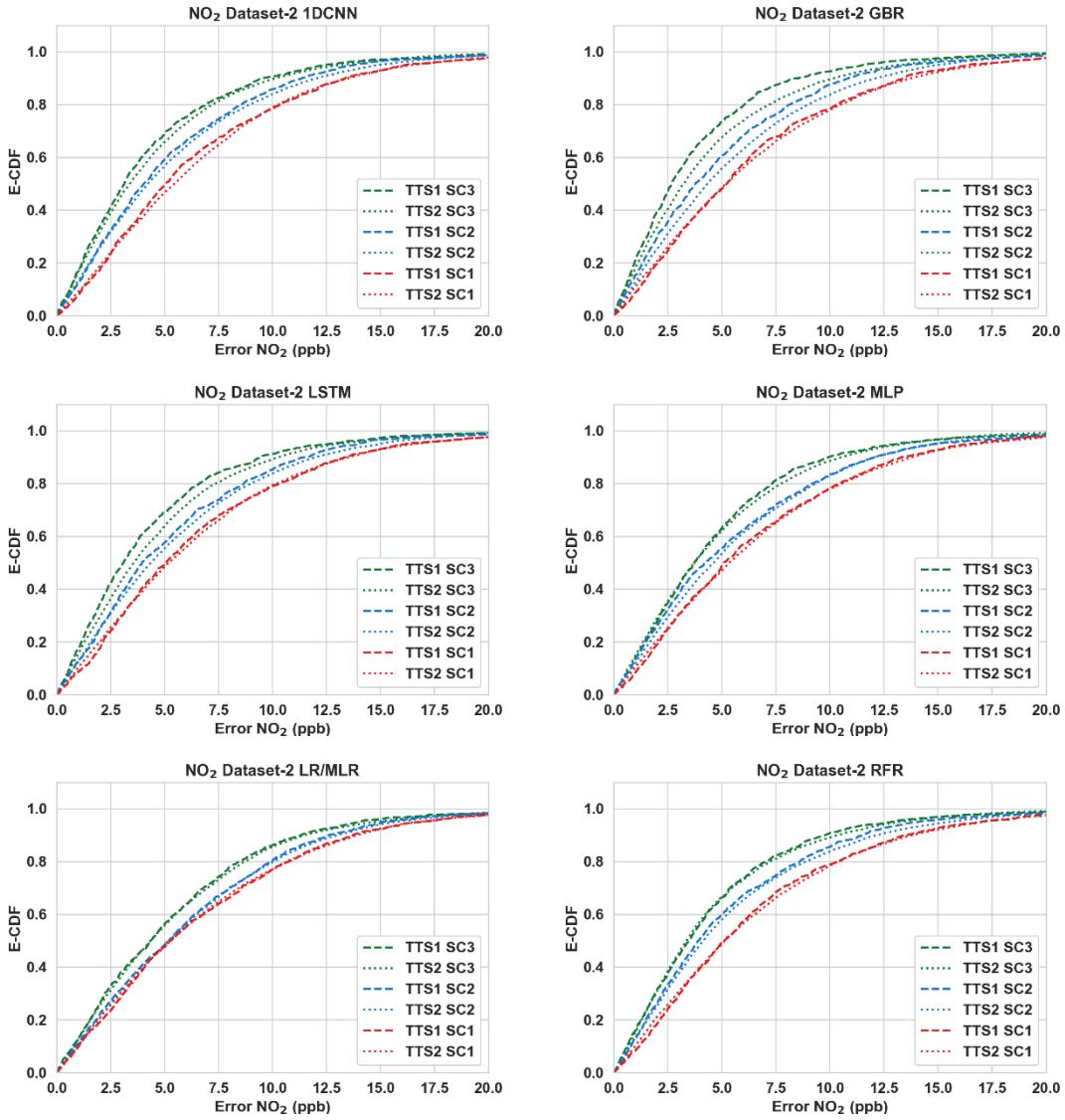


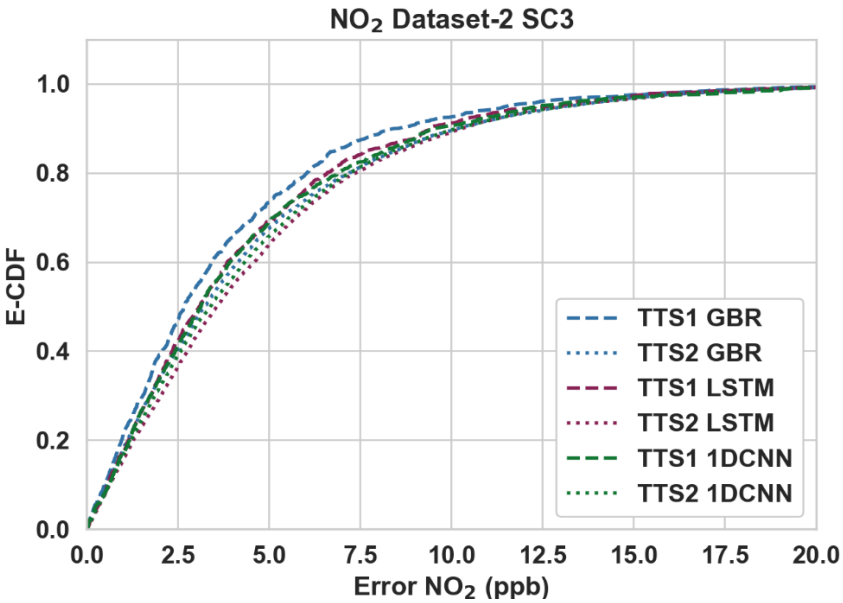
E-CDFs of CO Dataset 2 for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



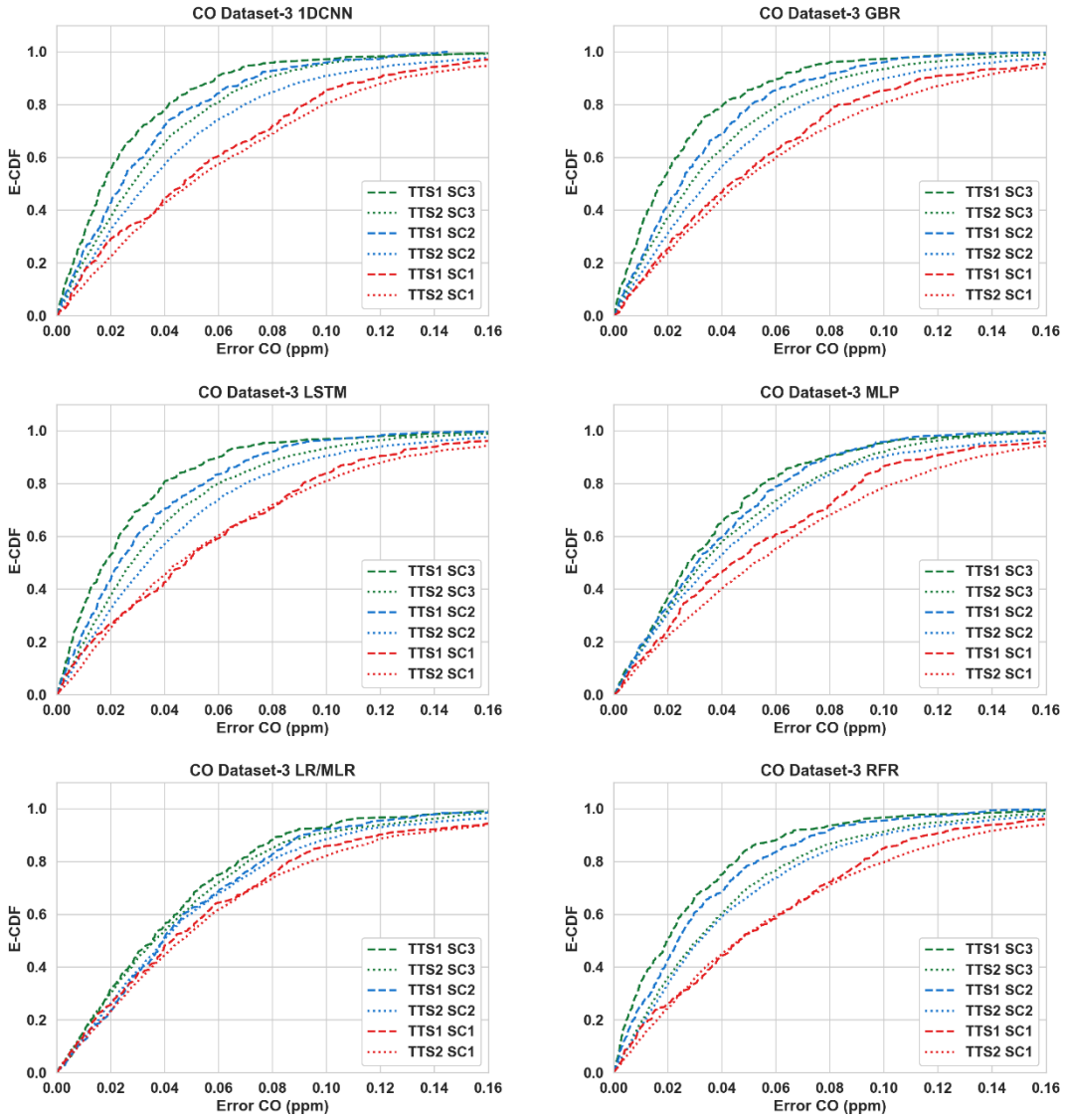


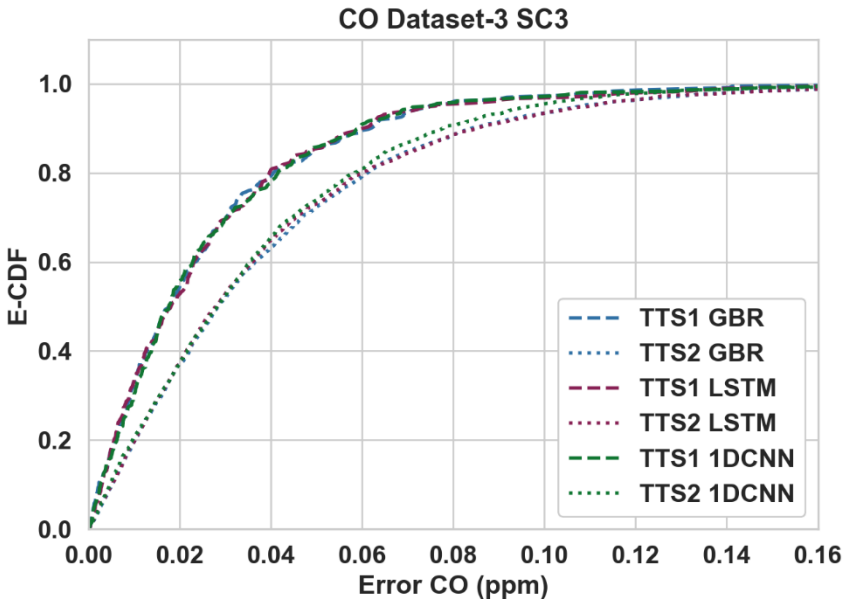
E-CDFs of NO_2 Dataset 2 for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



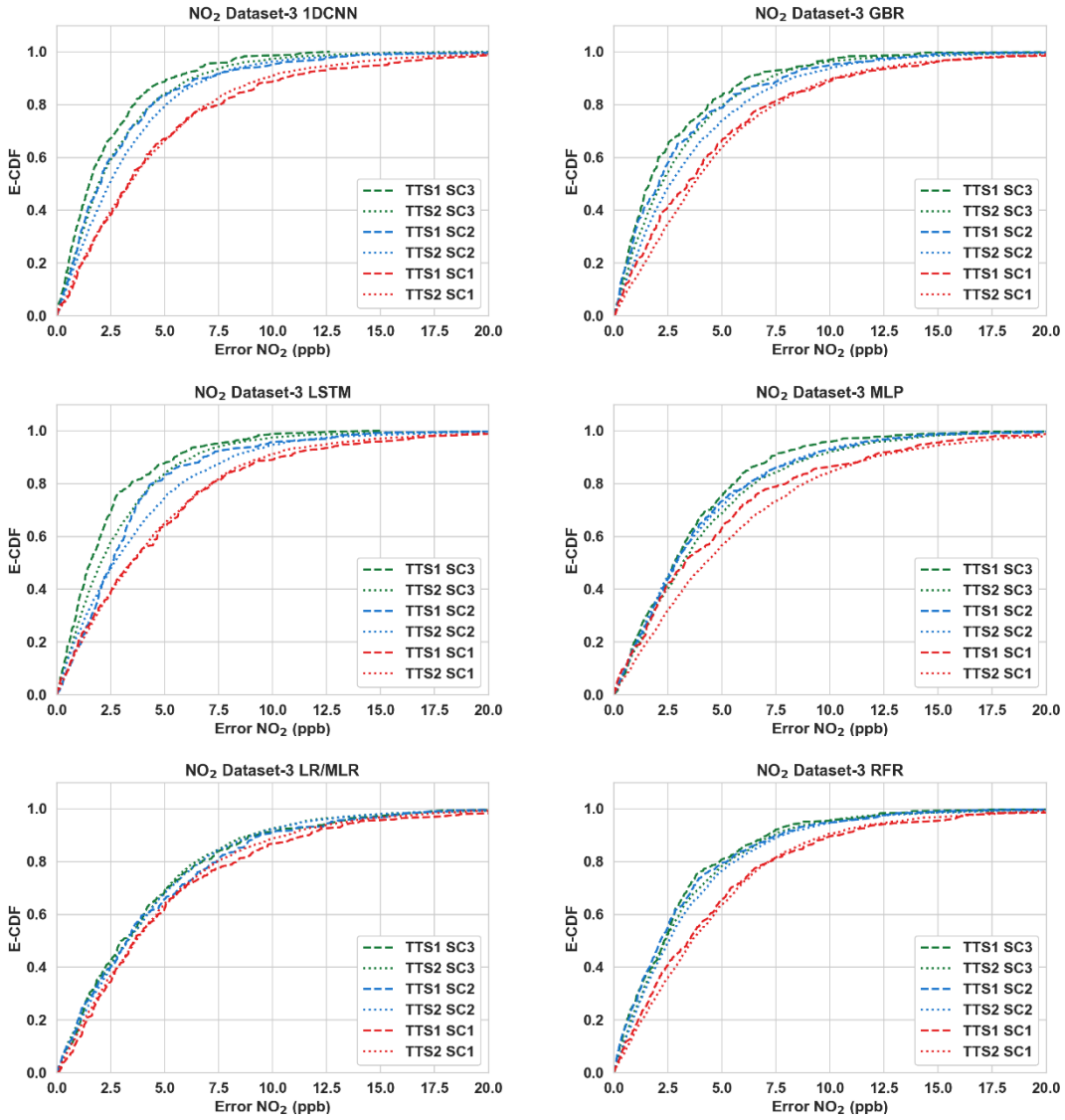


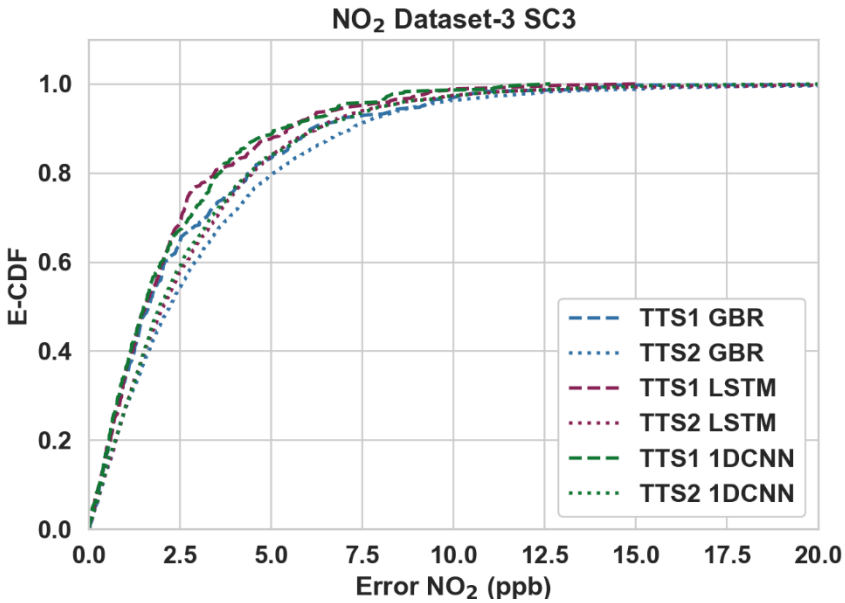
E-CDFs of CO Dataset 3 for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



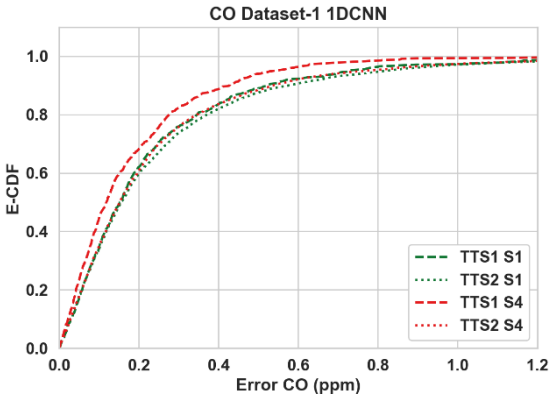


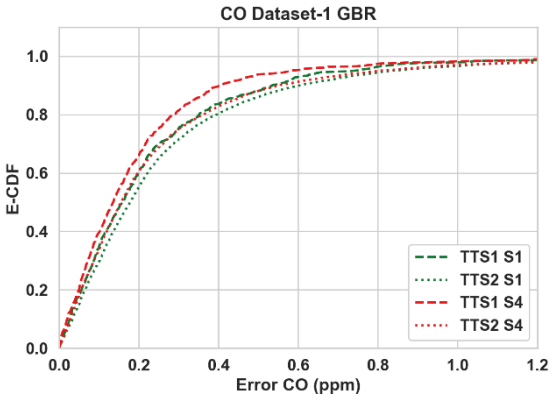
E-CDFs of NO_2 Dataset 3 for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)

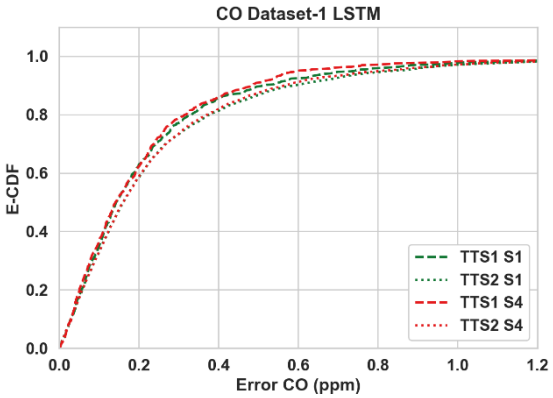


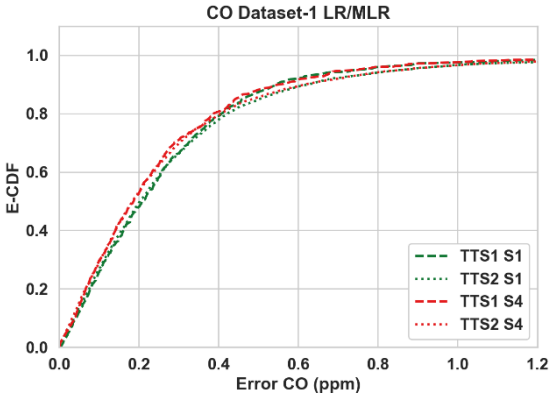


E-CDFs of CO Dataset 1 for different algorithms, between scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases), S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

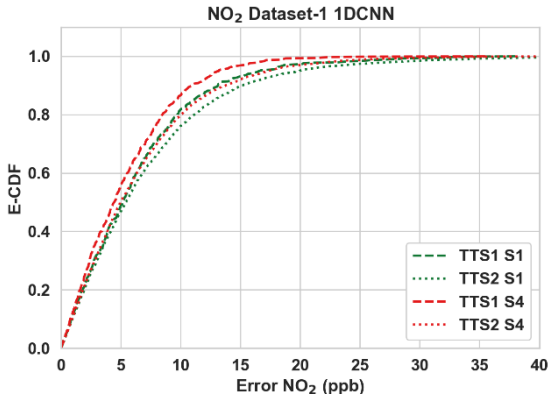


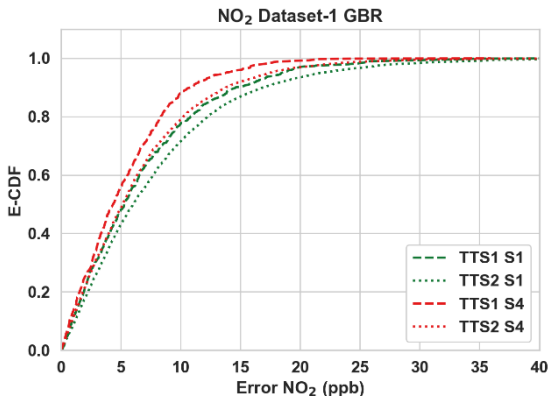


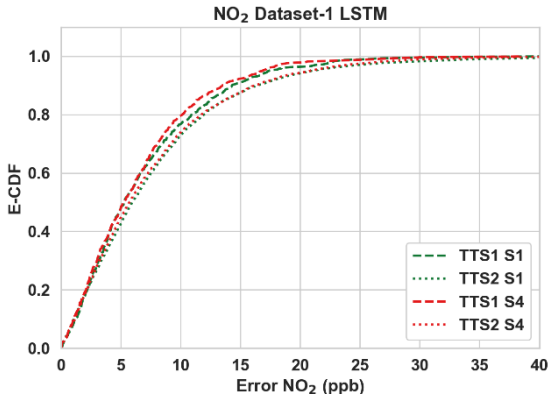


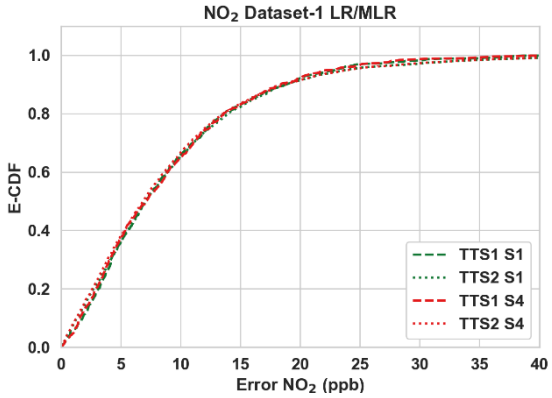


E-CDFs of NO_2 Dataset 1 for different algorithms, between scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases), S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

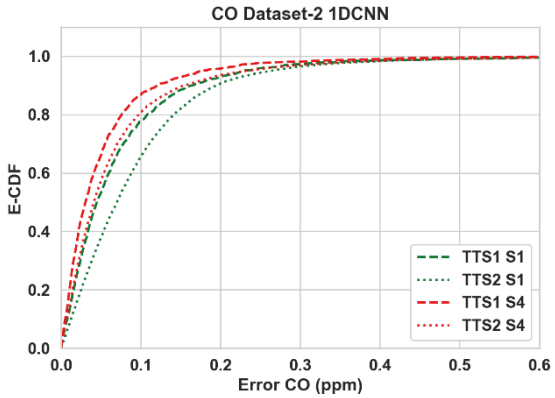


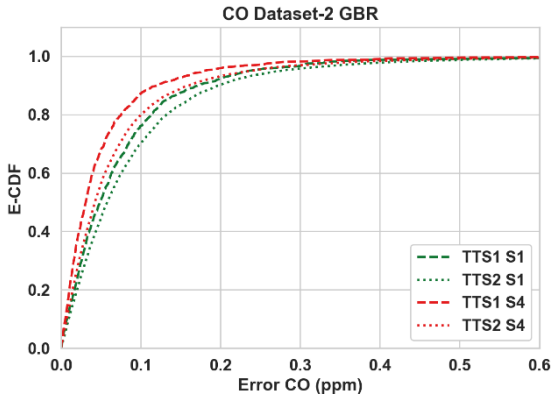


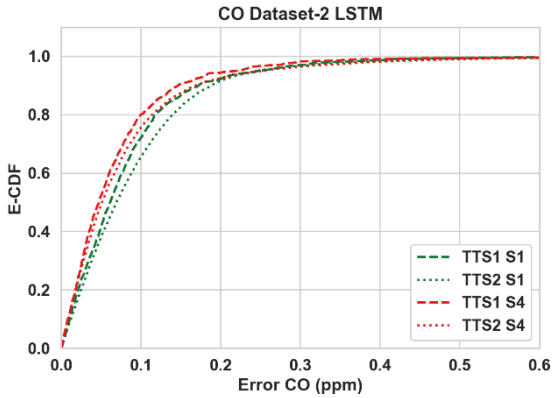


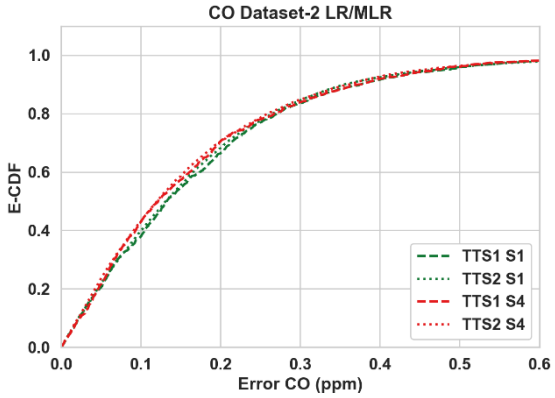


E-CDFs of CO Dataset 2 for different algorithms, between scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases), S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

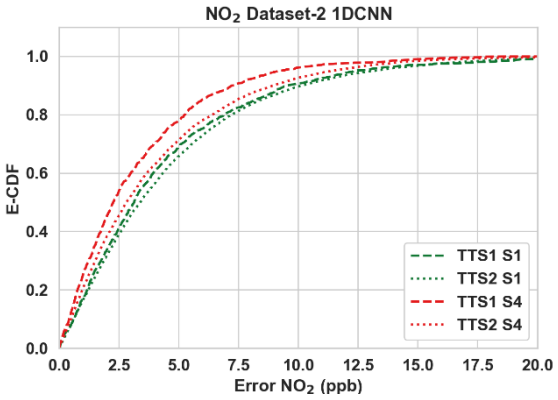


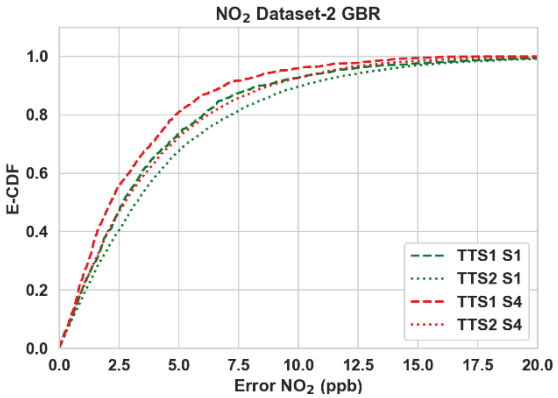


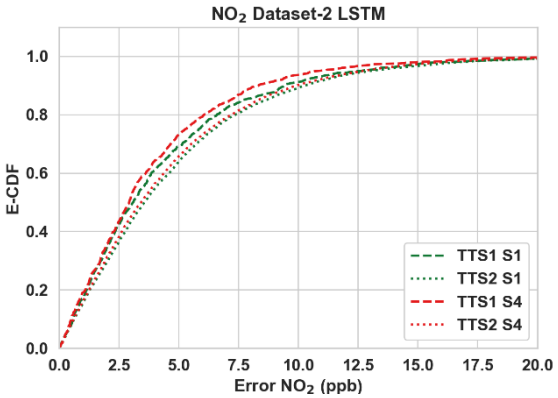


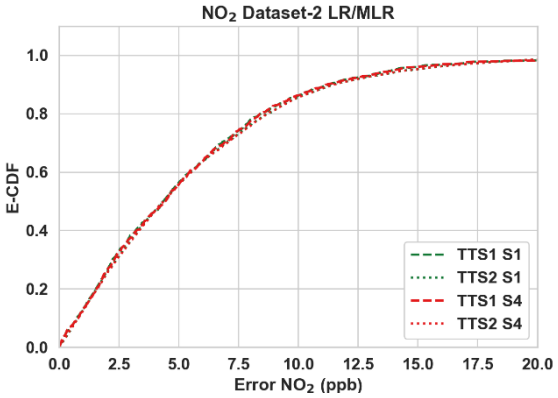


E-CDFs of NO_2 Dataset 2 for different algorithms, between scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases), S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

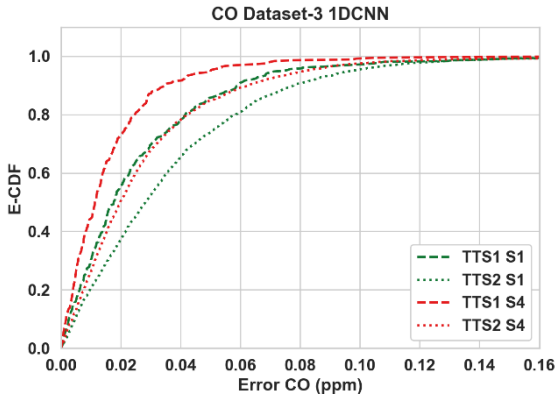


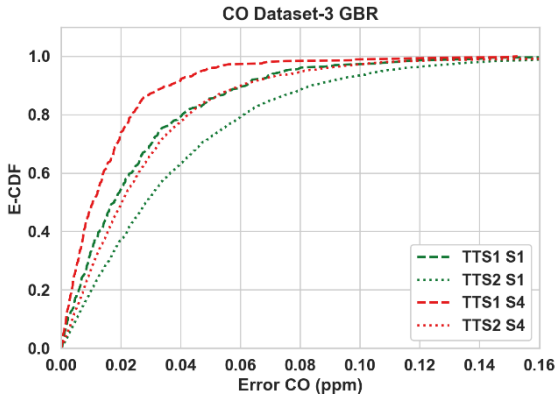


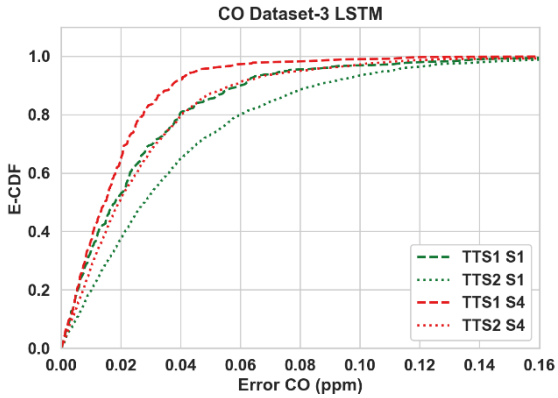


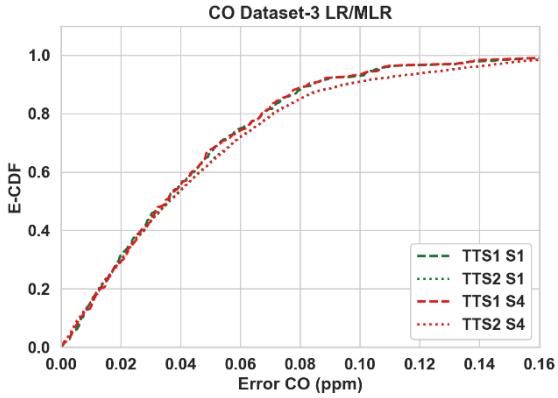


E-CDFs of CO Dataset 3 for different algorithms, between scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases), S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

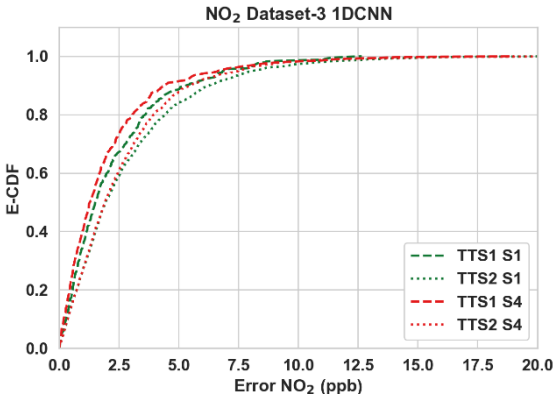


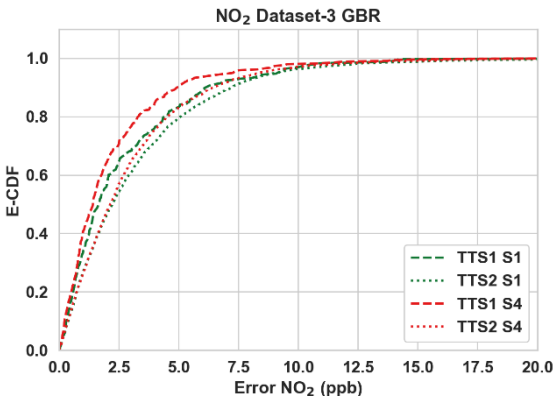


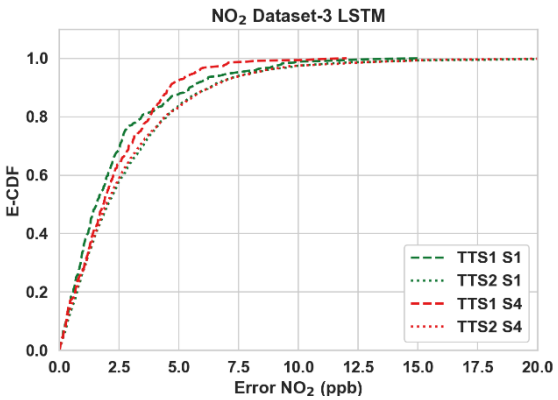


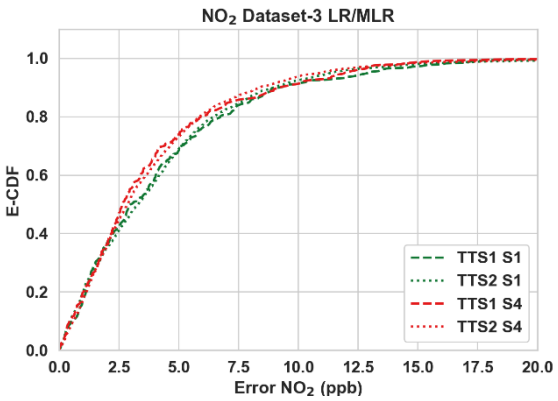


E-CDFs of NO_2 Dataset 3 for different algorithms, between scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases), S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

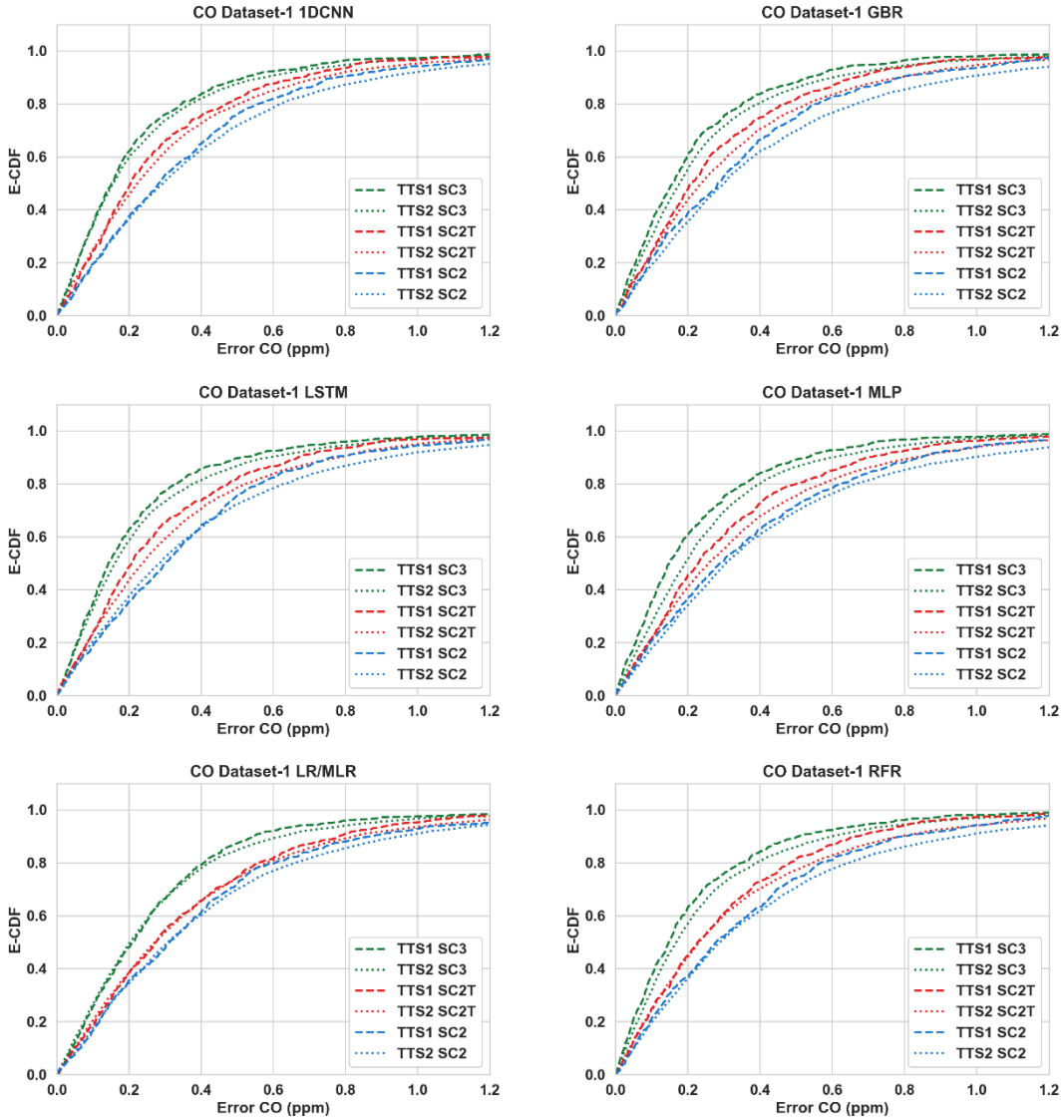




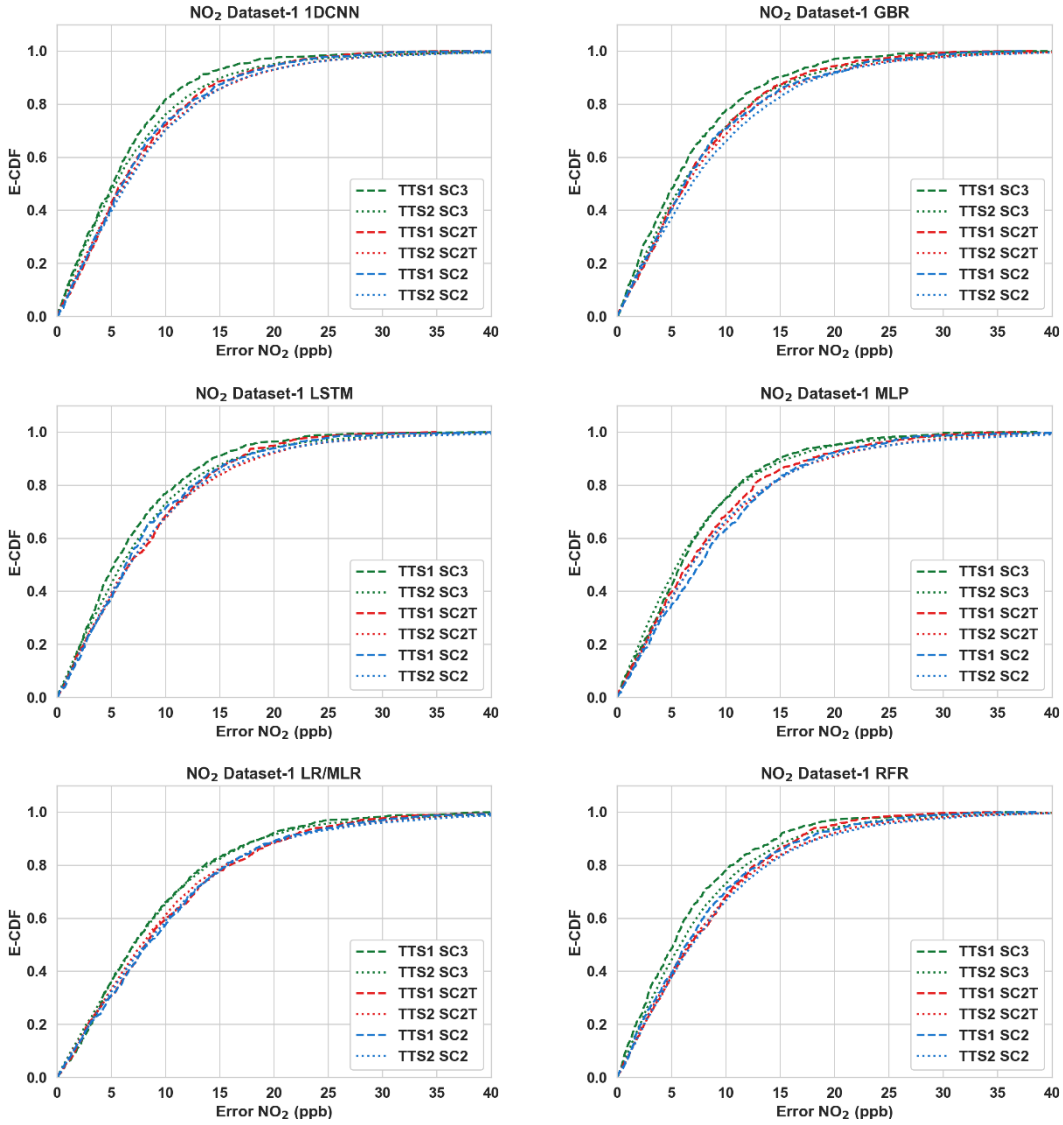




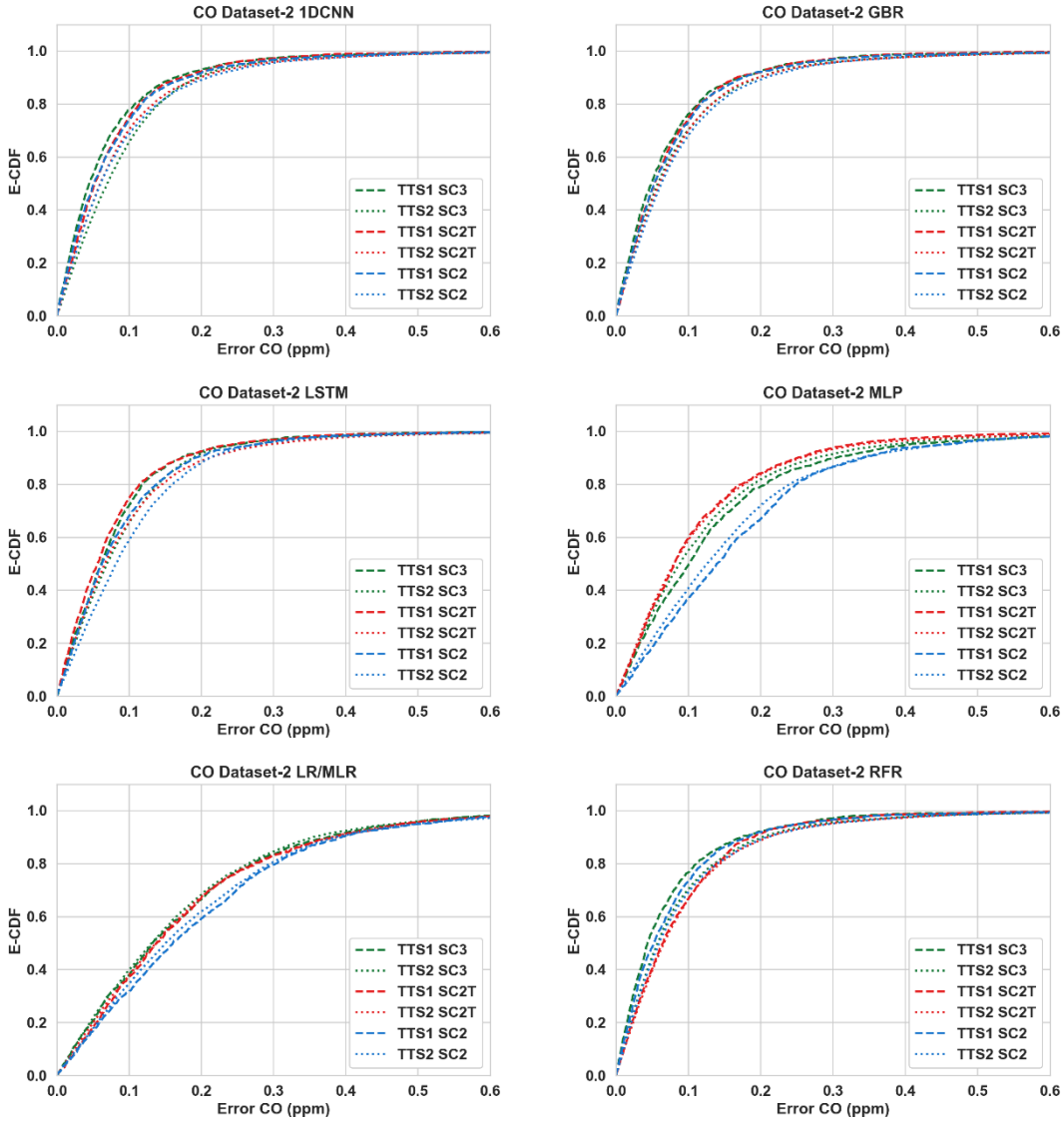
E-CDFs of CO Dataset 1 for different algorithms, between scenarios SC2 (raw LCS + Temp + Hum), SC2T (raw LCS + Temp + Hum+ N_{days} + Hour) and SC3 (same as S1 – raw LCS + Temp + Hum + other gases), and train test splits (TTS1 – 90:10, TTS2 – 20:80)



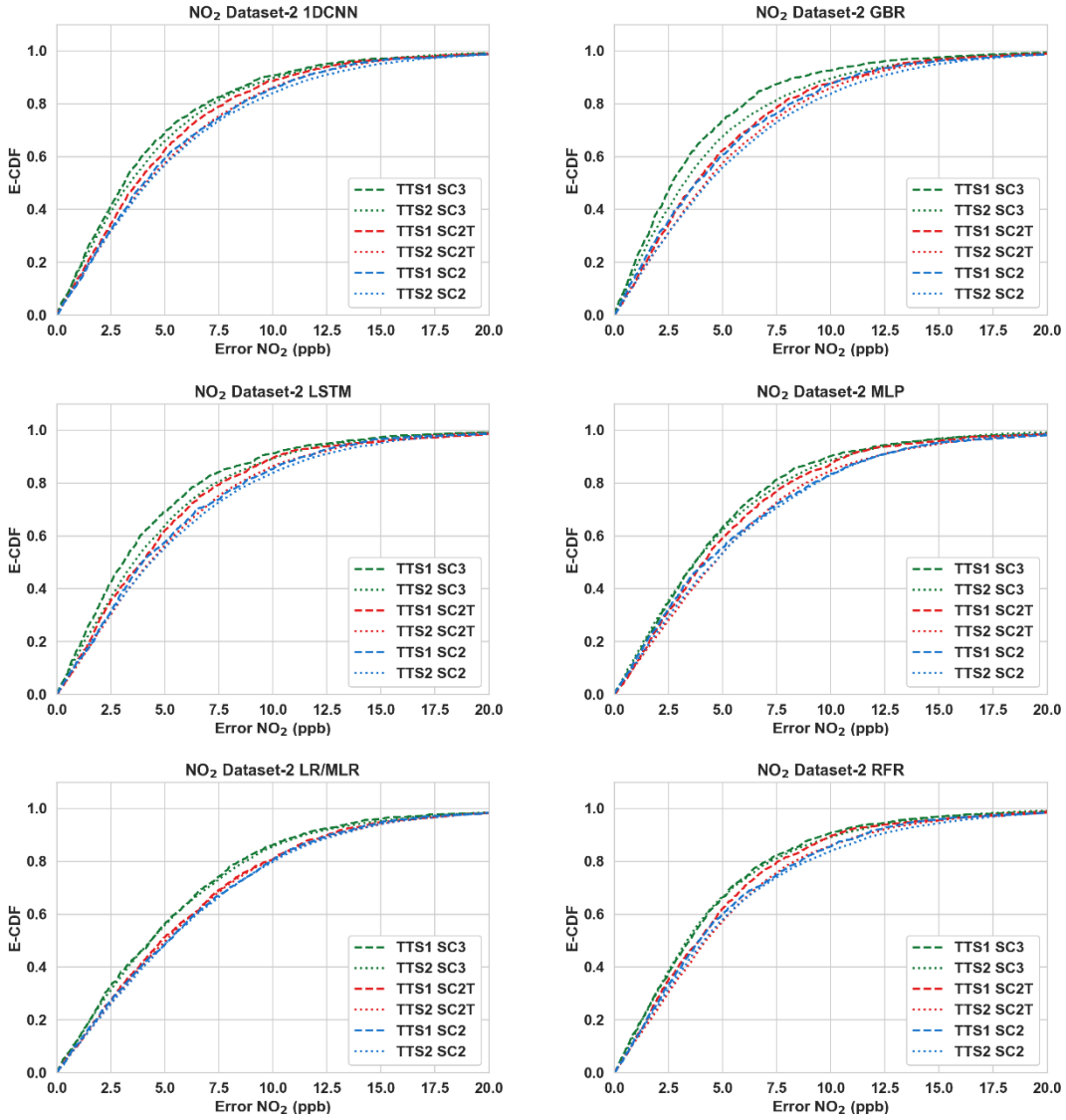
E-CDFs of NO_2 Dataset 1 for different algorithms, between scenarios SC2 (raw LCS + Temp + Hum), SC2T (raw LCS + Temp + Hum+ N_{days} + Hour) and SC3 (same as S1 – raw LCS + Temp + Hum + other gases), and train test splits (TTS1 – 90:10, TTS2 – 20:80)



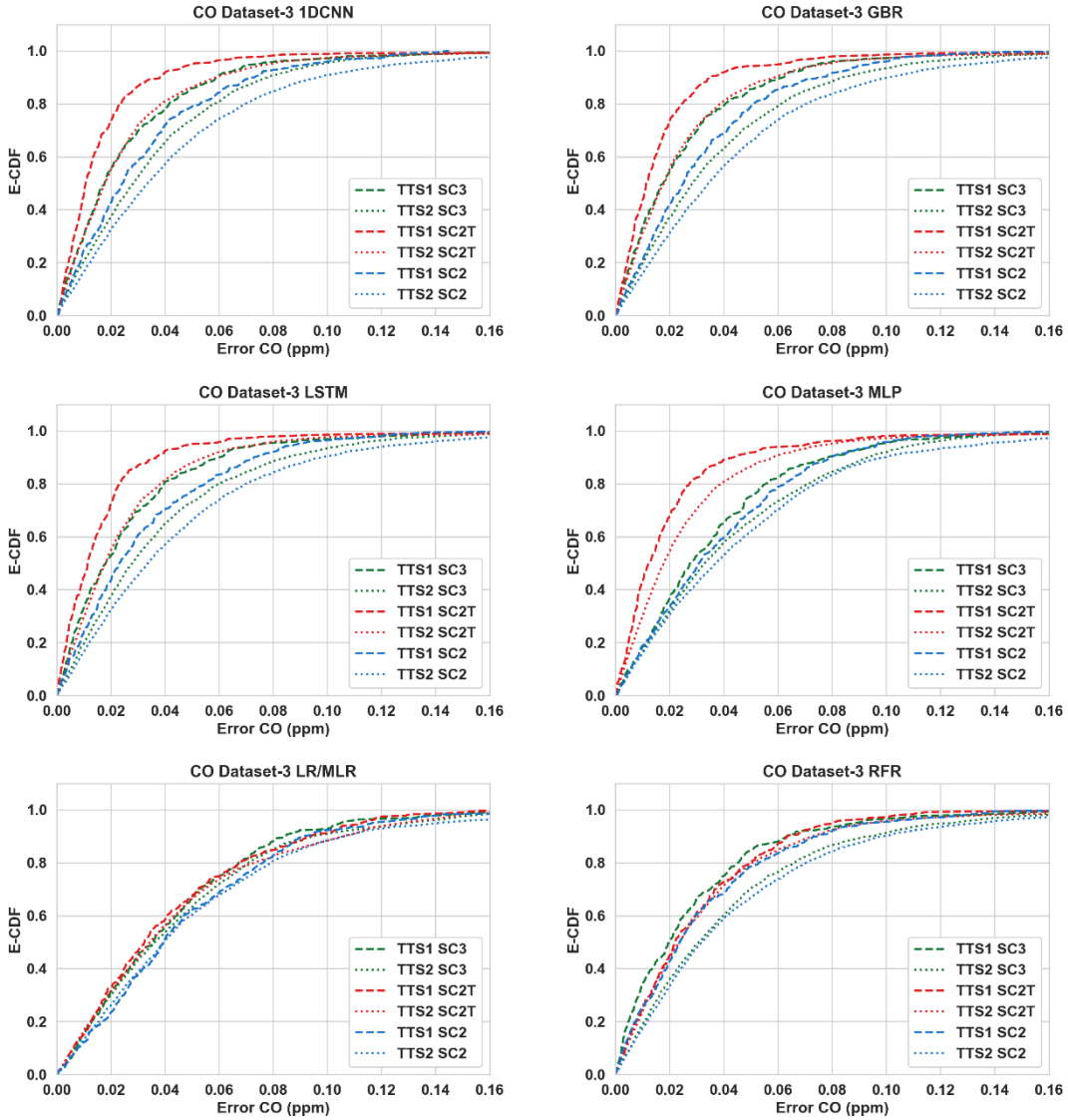
E-CDFs of CO Dataset 2 for different algorithms, between scenarios SC2 (raw LCS + Temp + Hum), SC2T (raw LCS + Temp + Hum+ N_{days} + Hour) and SC3 (same as S1 – raw LCS + Temp + Hum + other gases), and train test splits (TTS1 – 90:10, TTS2 – 20:80)



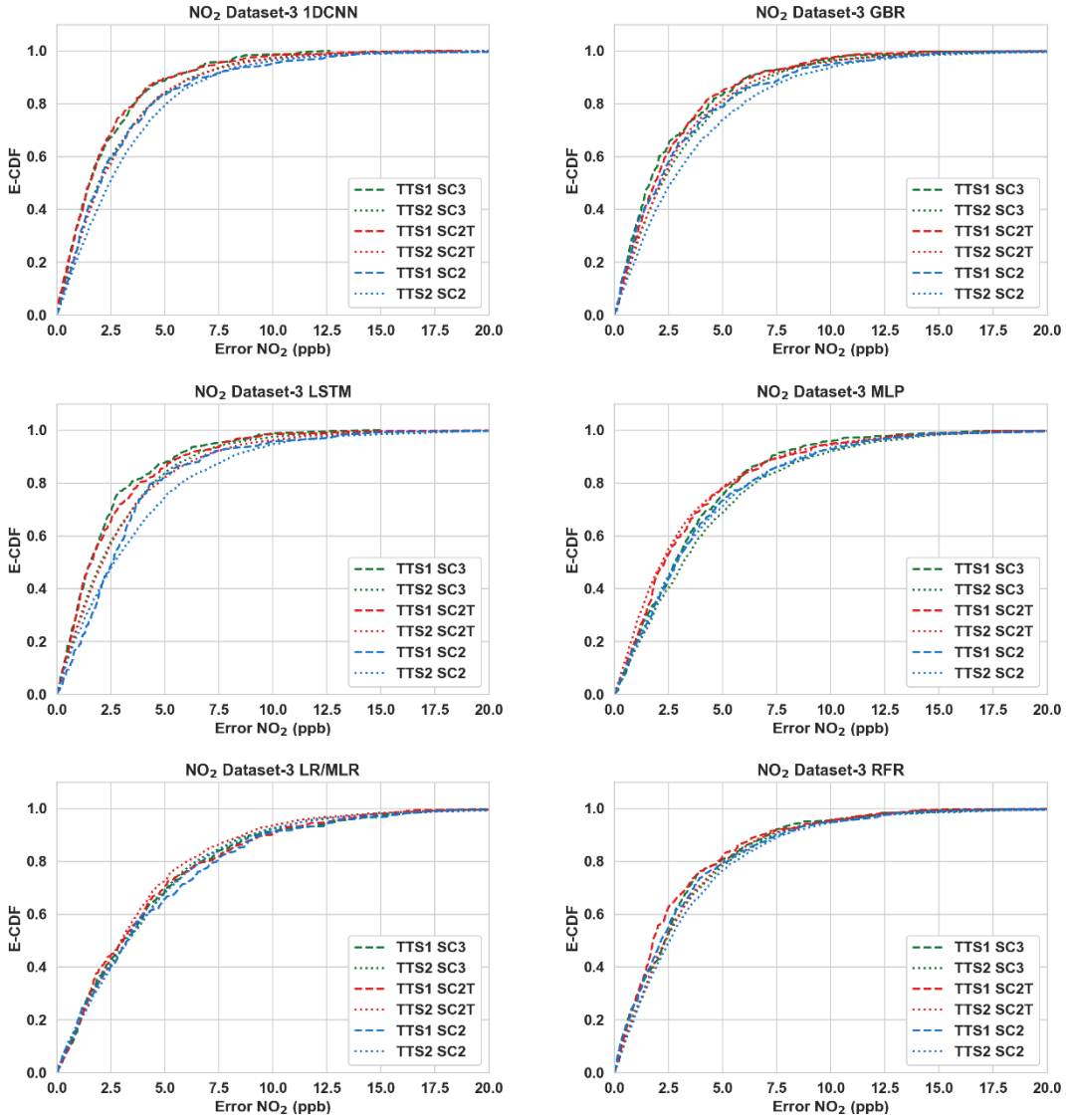
E-CDFs of NO_2 Dataset 2 for different algorithms, between scenarios SC2 (raw LCS + Temp + Hum), SC2T (raw LCS + Temp + Hum+ N_{days} + Hour) and SC3 (same as S1 – raw LCS + Temp + Hum + other gases), and train test splits (TTS1 – 90:10, TTS2 – 20:80)



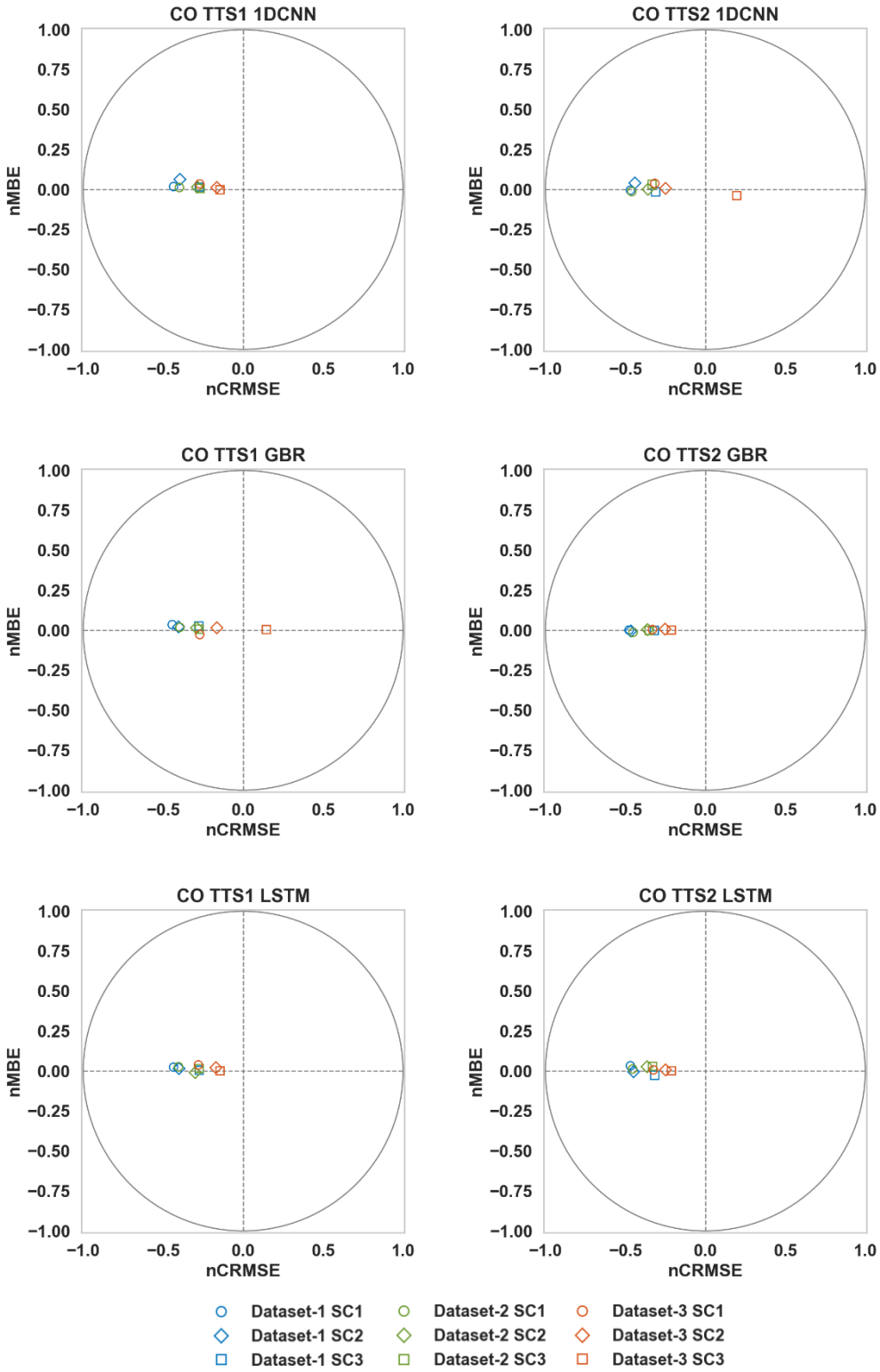
E-CDFs of CO Dataset 3 for different algorithms, between scenarios SC2 (raw LCS + Temp + Hum), SC2T (raw LCS + Temp + Hum+ N_{days} + Hour) and SC3 (same as S1 – raw LCS + Temp + Hum + other gases), and train test splits (TTS1 – 90:10, TTS2 – 20:80)



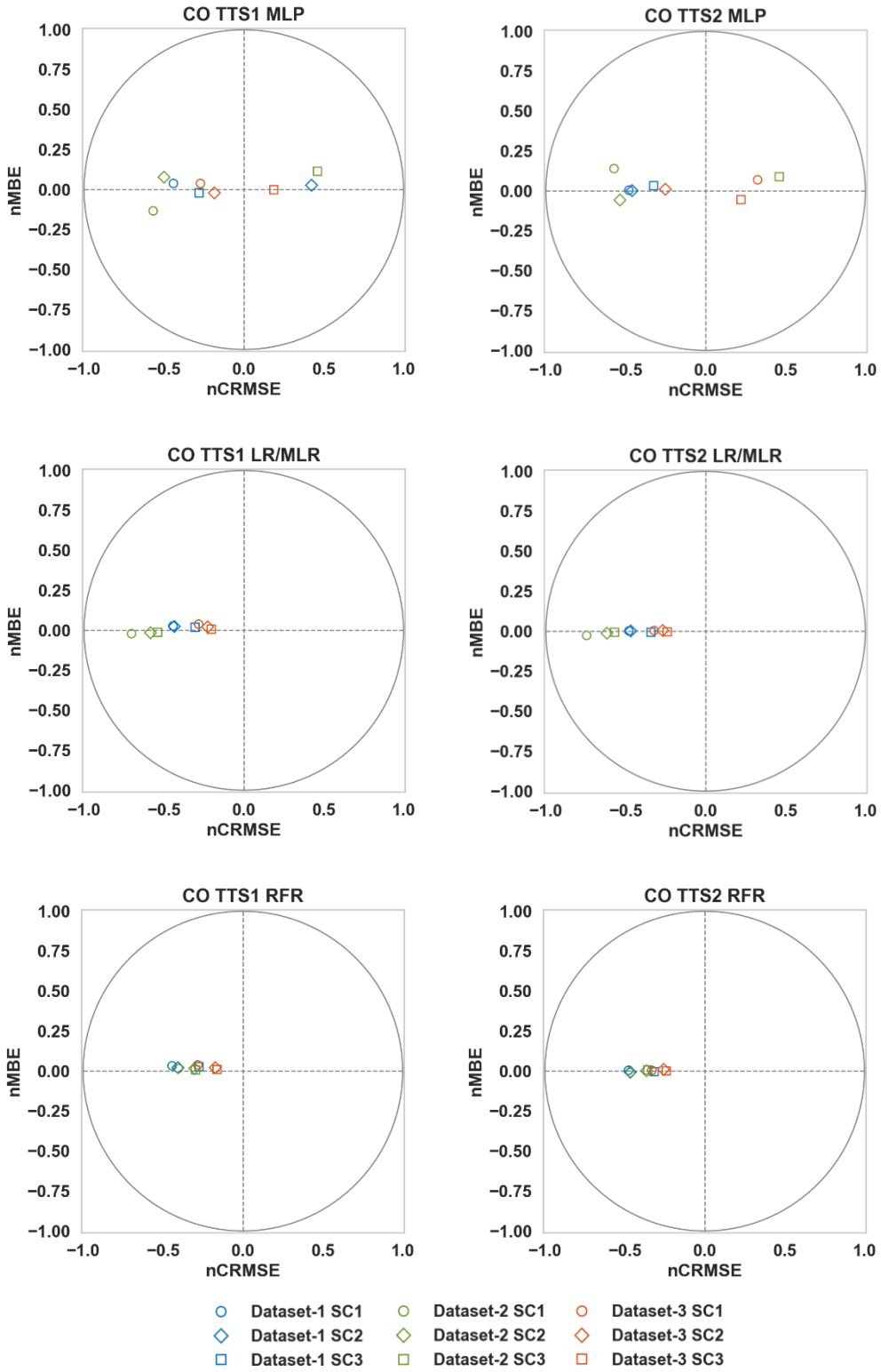
E-CDFs of NO_2 Dataset 3 for different algorithms, between scenarios SC2 (raw LCS + Temp + Hum), SC2T (raw LCS + Temp + Hum+ N_{days} + Hour) and SC3 (same as S1 – raw LCS + Temp + Hum + other gases), and train test splits (TTS1 – 90:10, TTS2 – 20:80)



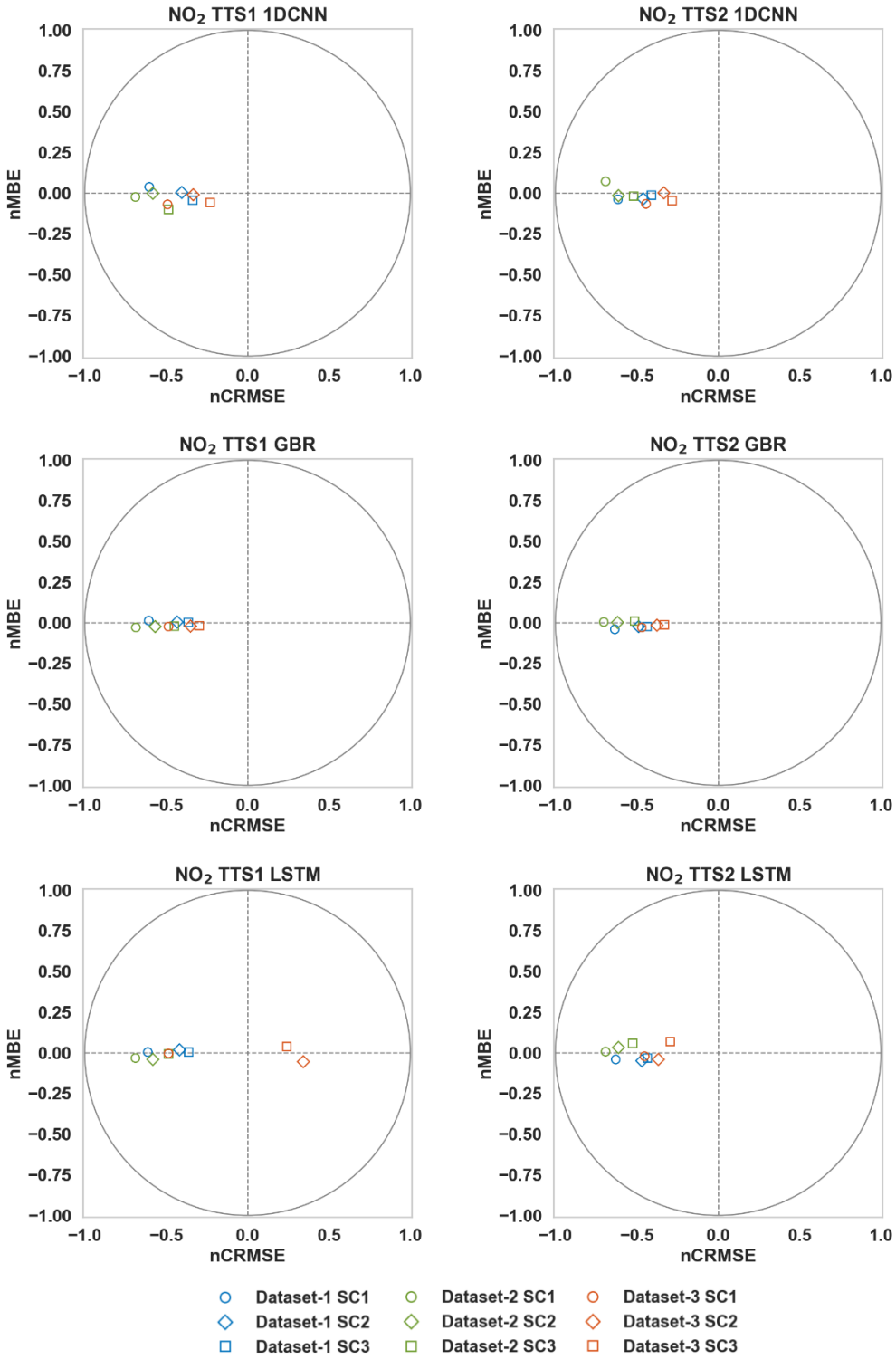
Target diagrams of CO from all three datasets for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



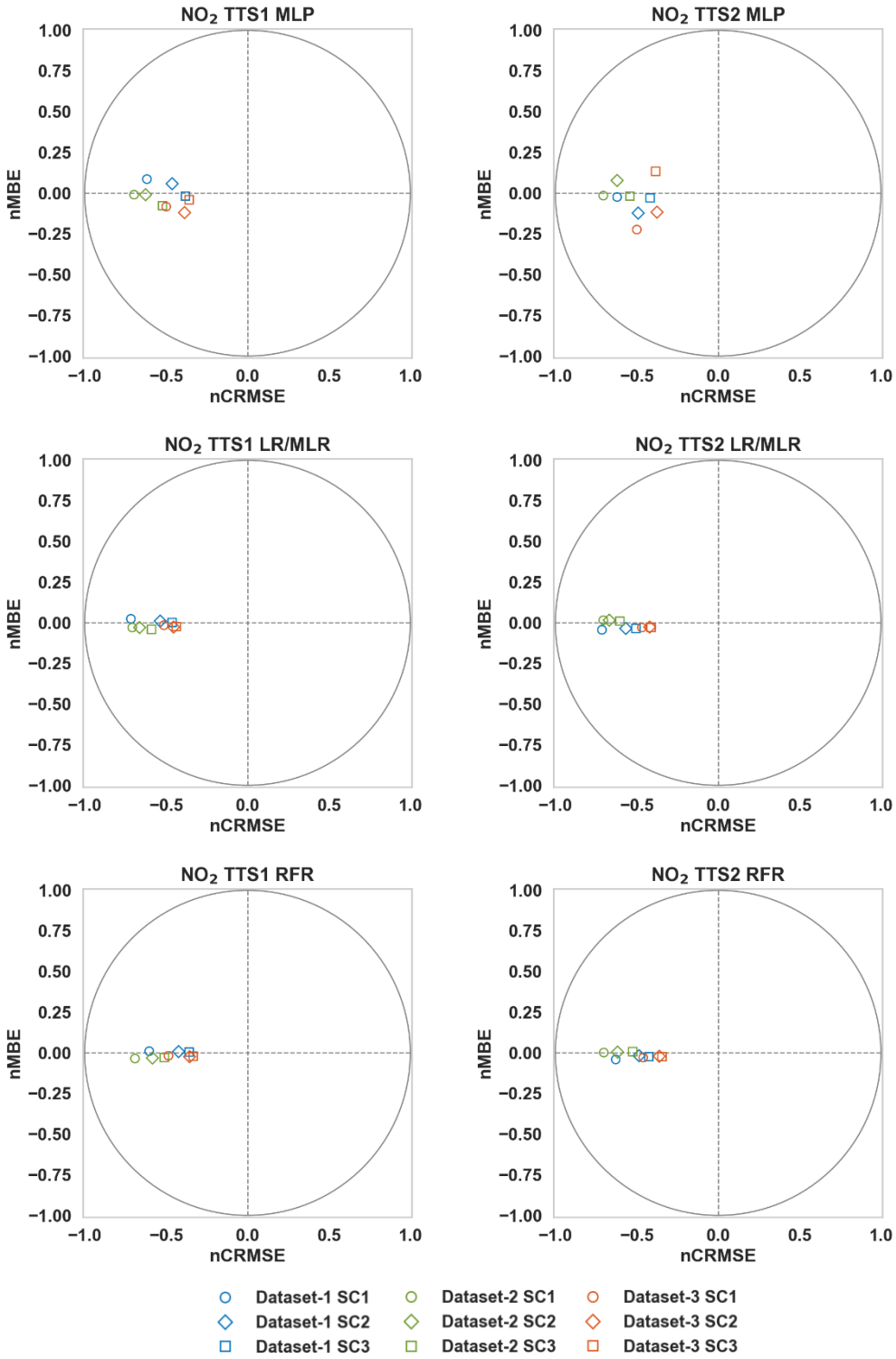
Target diagrams of CO from all three datasets for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



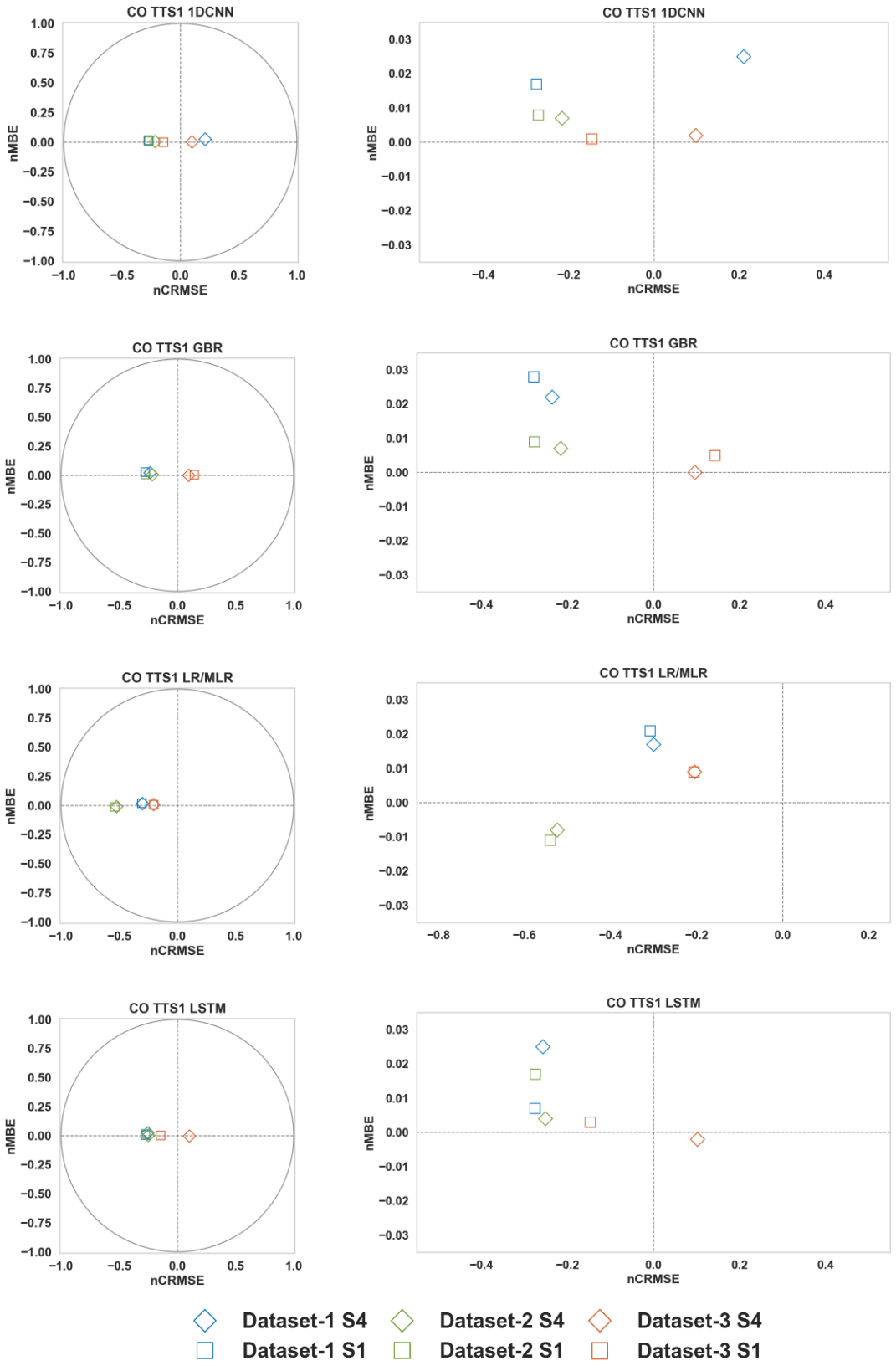
Target diagrams of NO_2 from all three datasets for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)

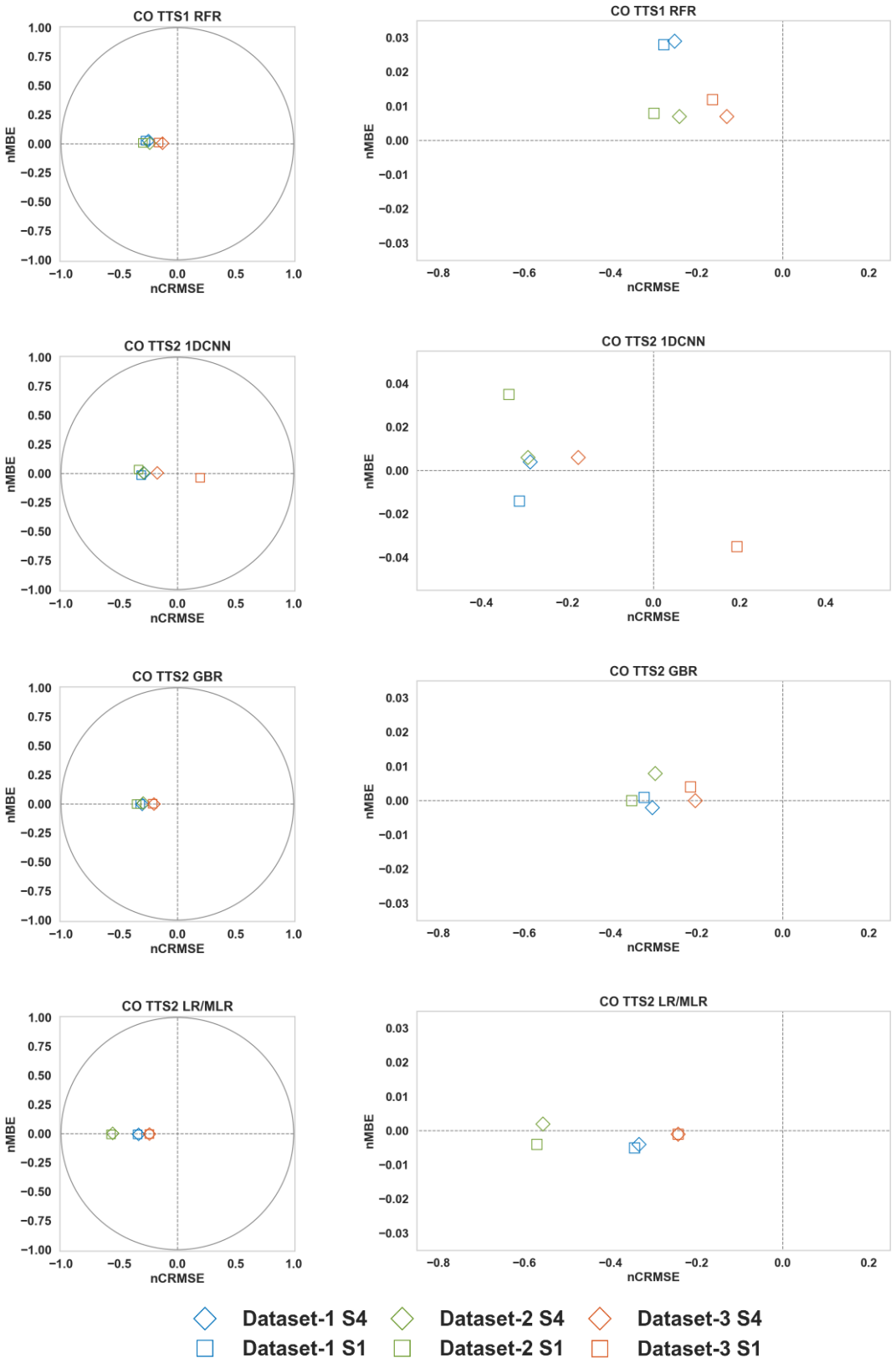


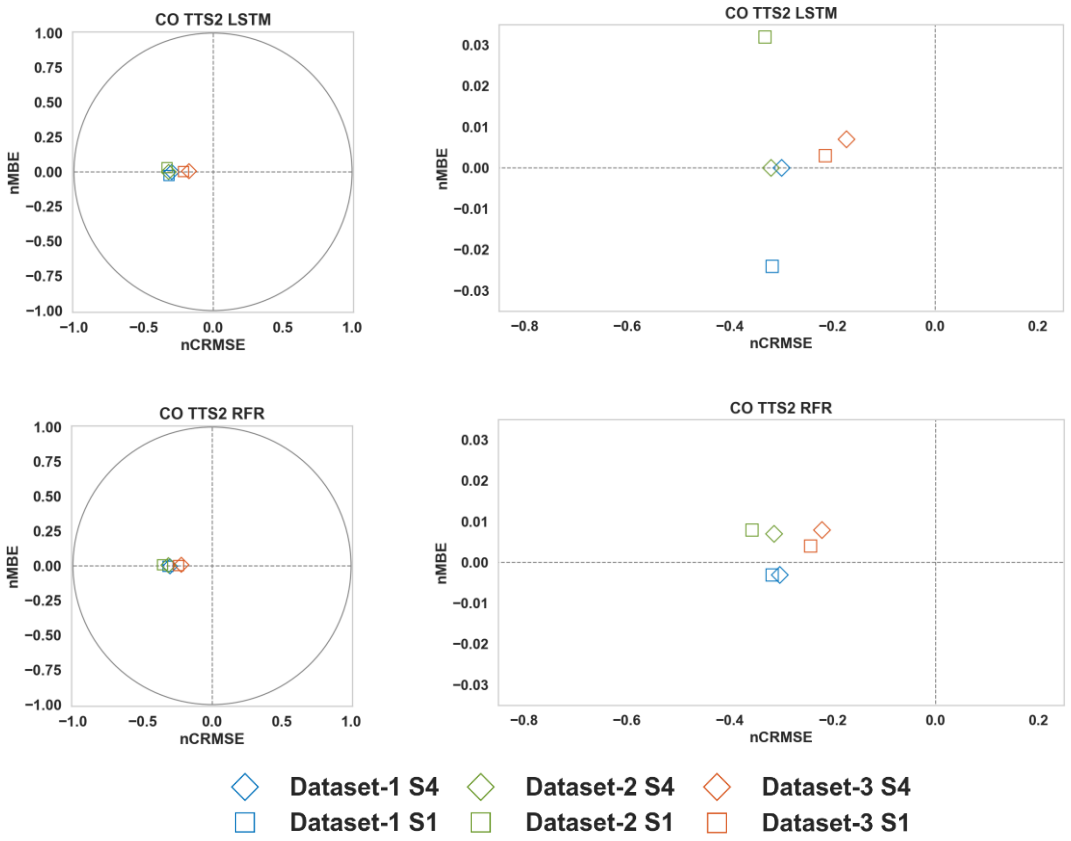
Target diagrams of NO_2 from all three datasets for different algorithms, scenarios (SC1 – raw LCS, SC2 – raw LCS + Temp + Hum, SC3 – raw LCS + Temp + Hum + other gases) and train test splits (TTS1 – 90:10, TTS2 – 20:80)



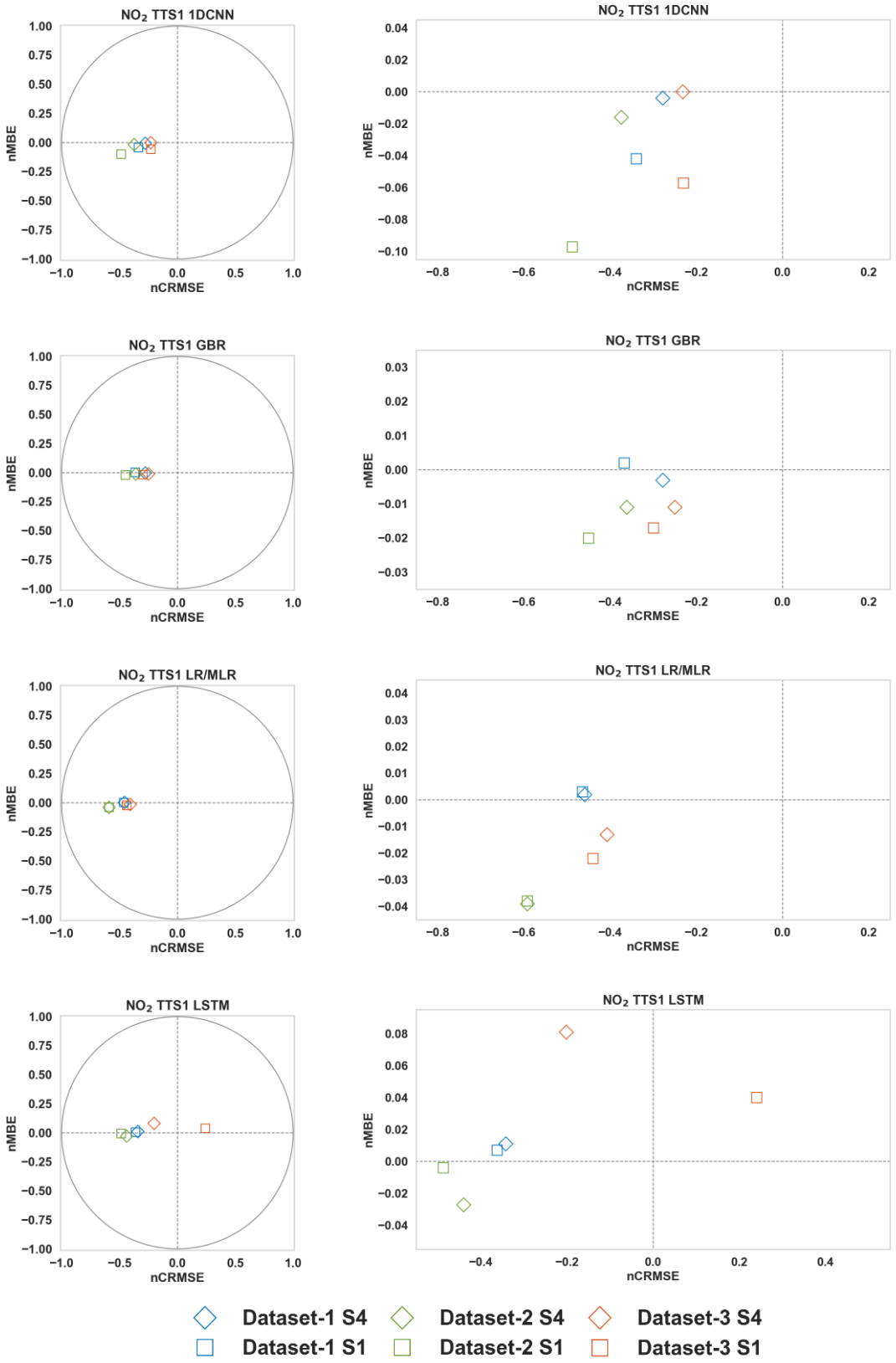
Target diagrams of CO from all three datasets for different algorithms, scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases, S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

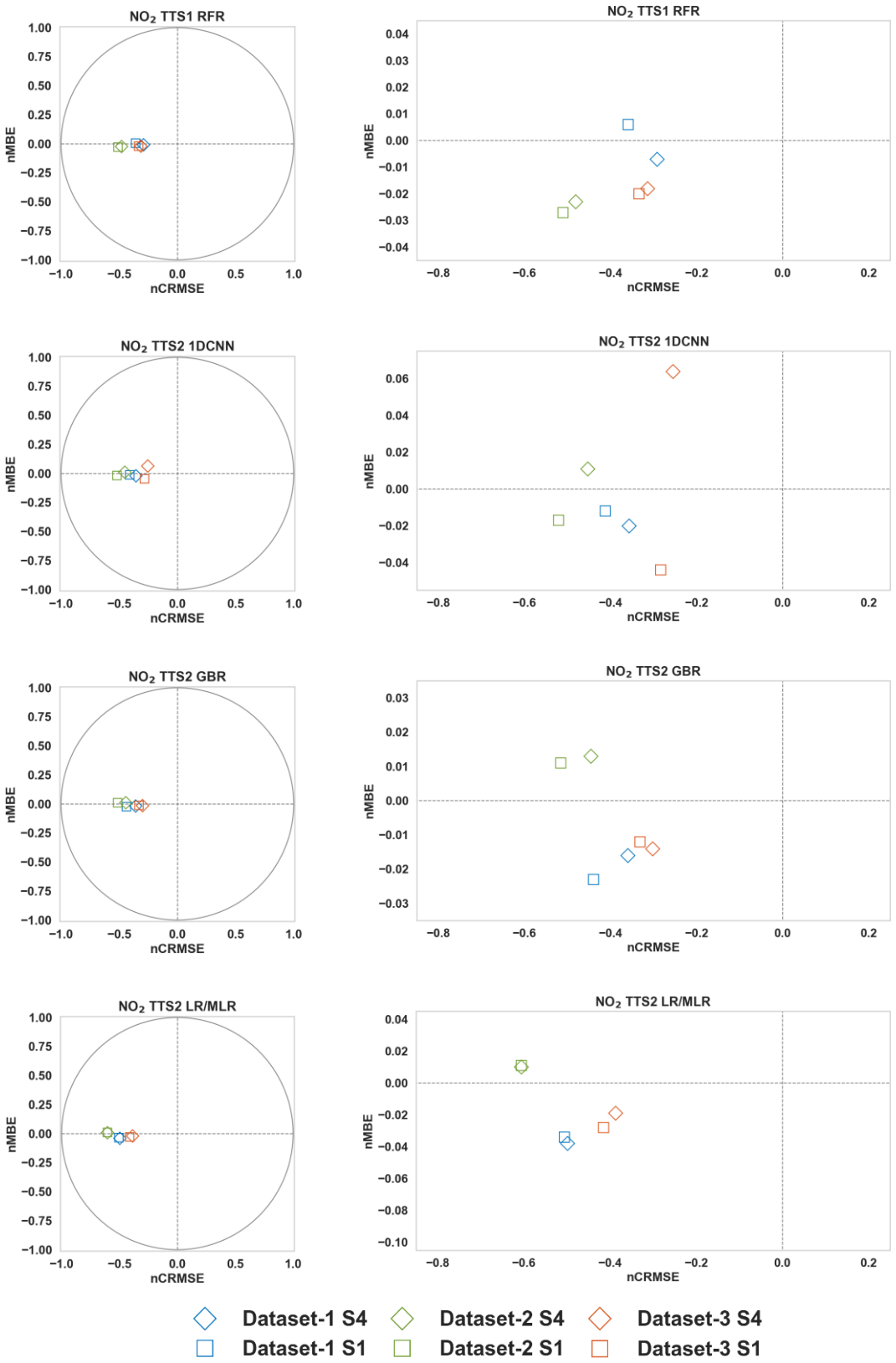


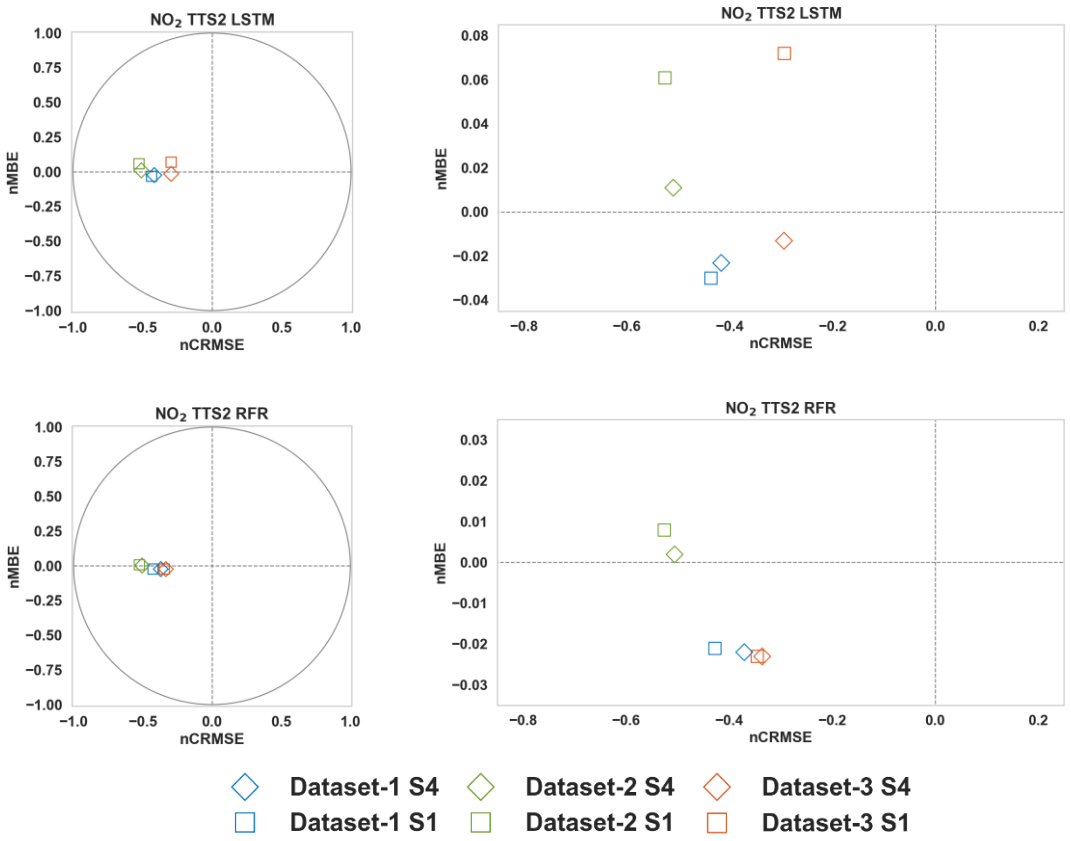

CO TTS1 RFR
CO TTS1 RFR
CO TTS2 1DCNN
CO TTS2 1DCNN
CO TTS2 GBR
CO TTS2 GBR
CO TTS2 LR/MLR
CO TTS2 LR/MLR
nMBE
nCRMSE
Dataset-1 S4
Dataset-2 S4
Dataset-3 S4
Dataset-1 S1
Dataset-2 S1
Dataset-3 S1


CO TTS2 LSTM
nMBE
nCRMSE
CO TTS2 LSTM
nMBE
nCRMSE
CO TTS2 RFR
nMBE
nCRMSE
CO TTS2 RFR
nMBE
nCRMSE
Dataset-1 S4
Dataset-2 S4
Dataset-3 S4
Dataset-1 S1
Dataset-2 S1
Dataset-3 S1

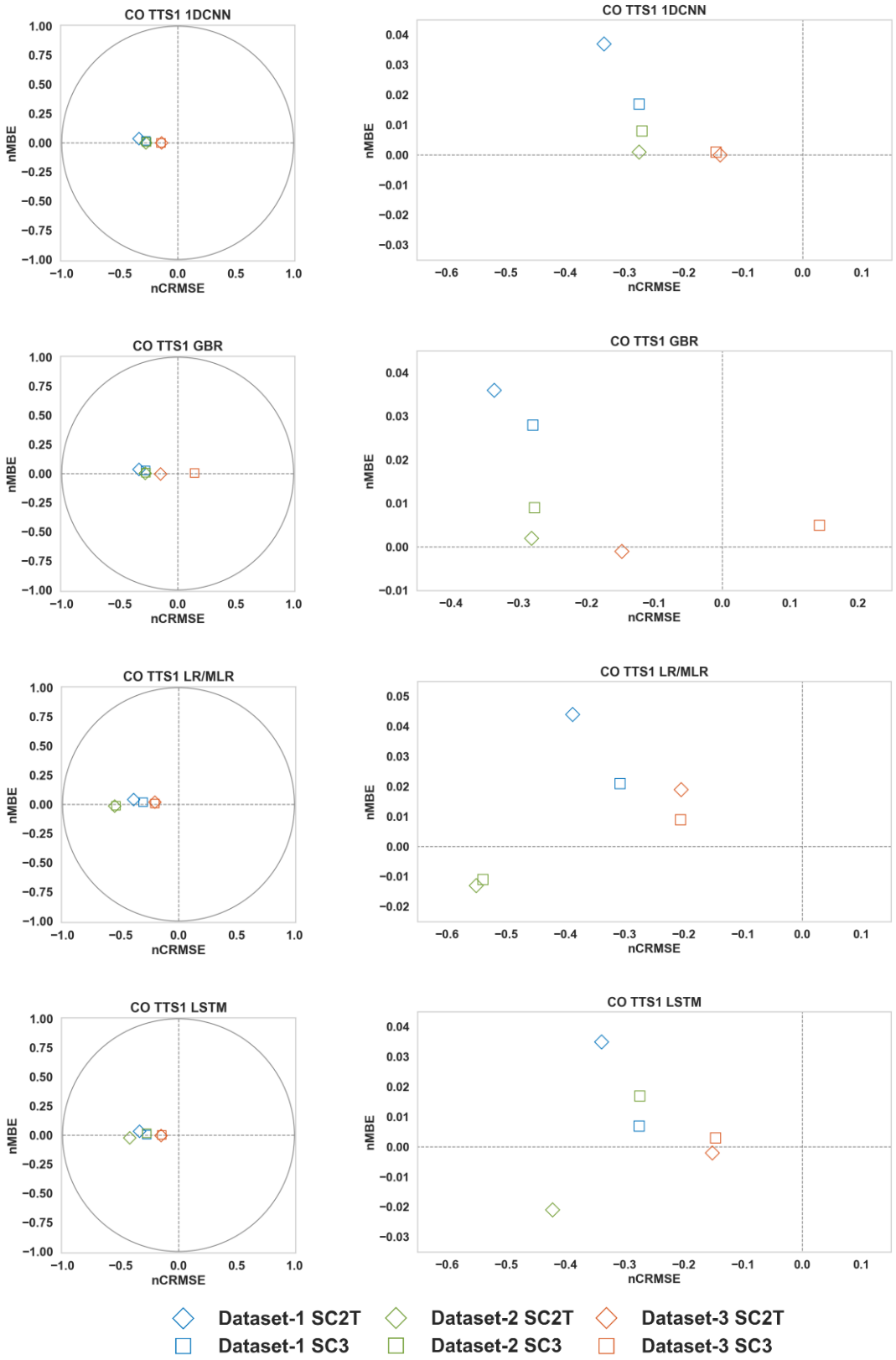
Target diagrams of NO_2 from all three datasets for different algorithms, scenarios: S1 (same as SC3) – raw LCS + Temp + Hum + other gases, S4 – raw LCS + Temp + Hum + other gases + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

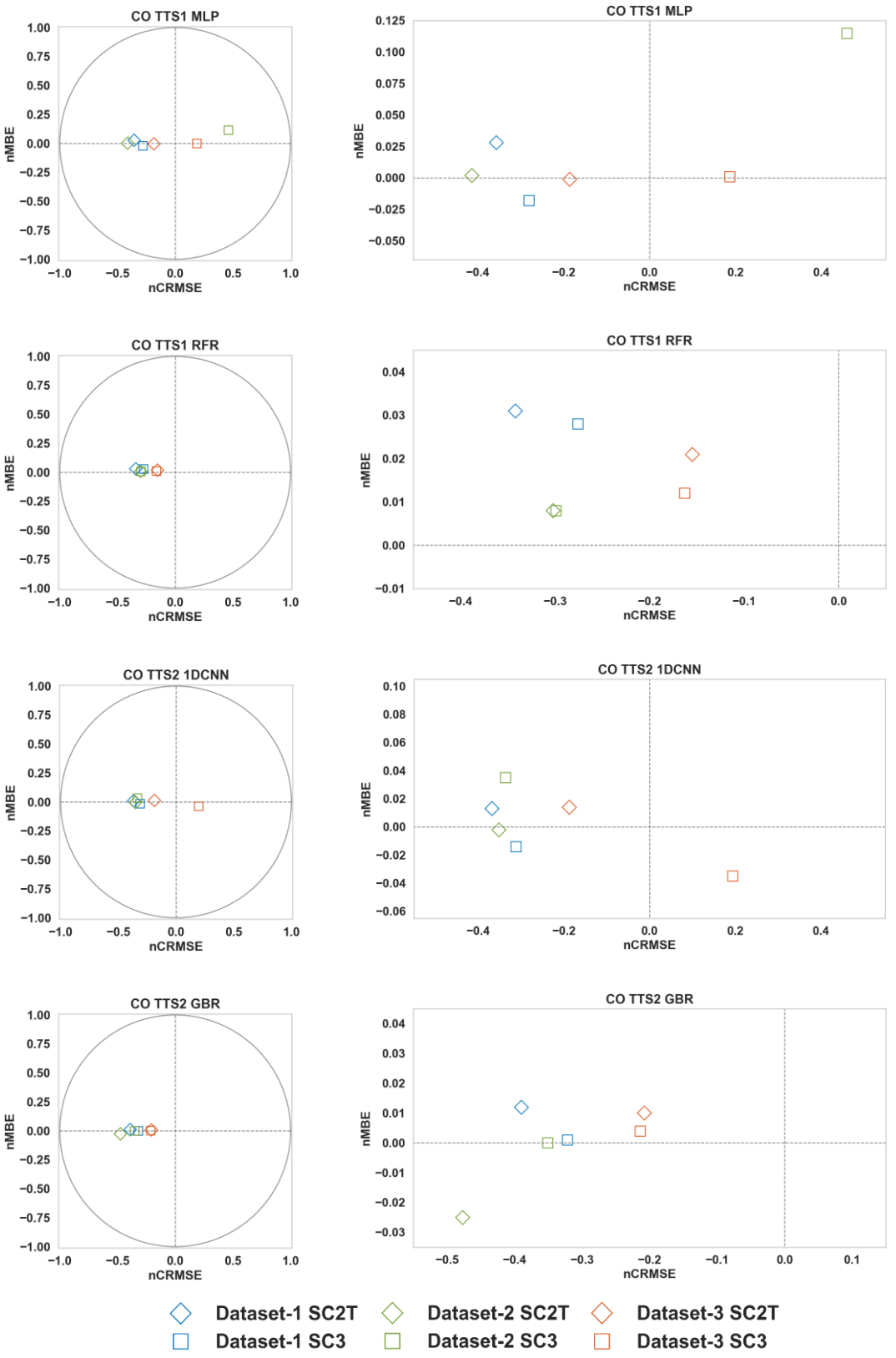


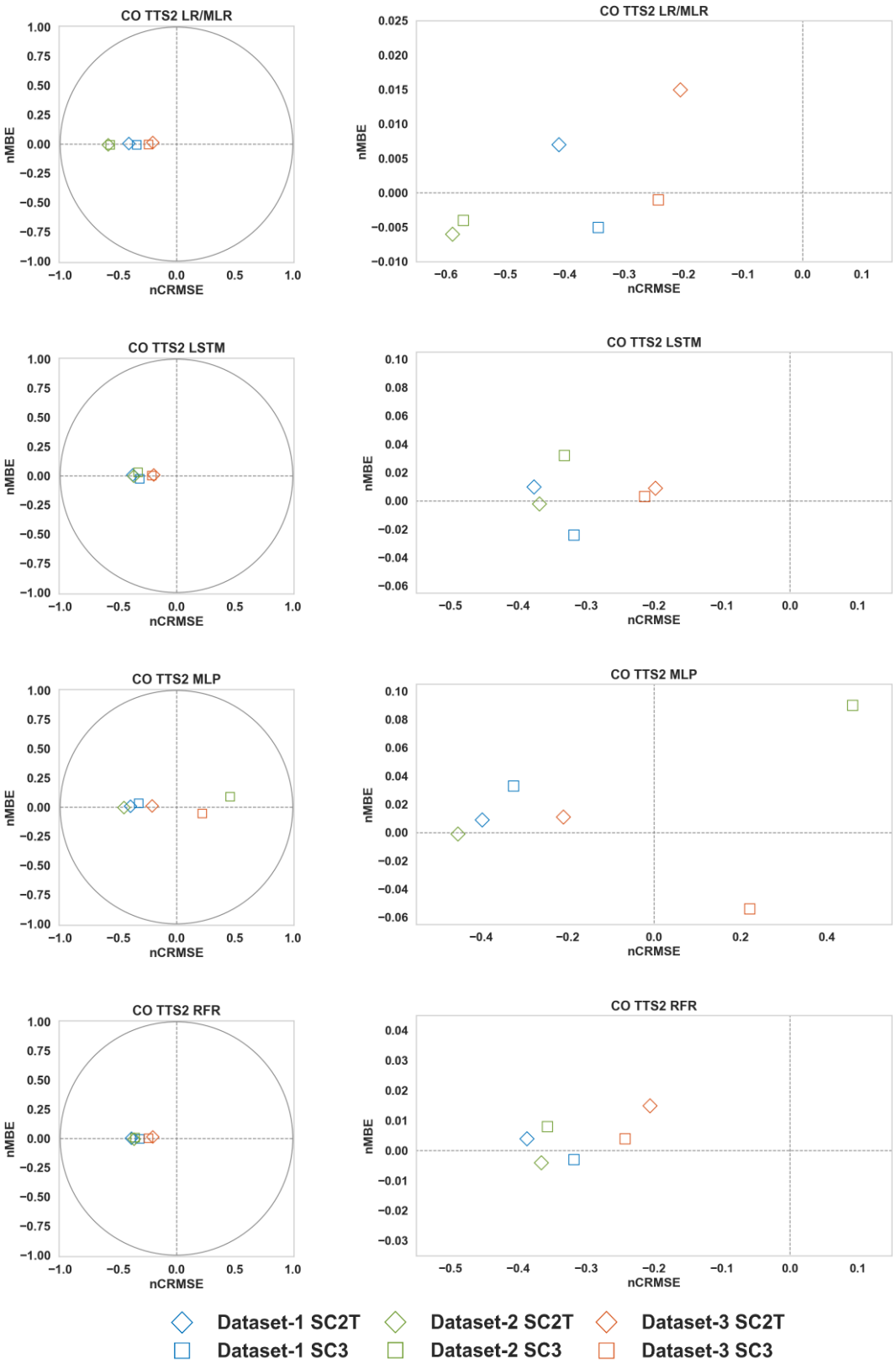

NO2 TTS1 RFR
NO2 TTS1 RFR
NO2 TTS2 1DCNN
NO2 TTS2 1DCNN
NO2 TTS2 GBR
NO2 TTS2 GBR
NO2 TTS2 LR/MLR
NO2 TTS2 LR/MLR
nMBE
nCRMSE
Dataset-1 S4
Dataset-2 S4
Dataset-3 S4
Dataset-1 S1
Dataset-2 S1
Dataset-3 S1


NO2 TTS2 LSTM
nMBE
nCRMSE
NO2 TTS2 LSTM
nMBE
nCRMSE
NO2 TTS2 RFR
nMBE
nCRMSE
NO2 TTS2 RFR
nMBE
nCRMSE
Dataset-1 S4
Dataset-2 S4
Dataset-3 S4
Dataset-1 S1
Dataset-2 S1
Dataset-3 S1

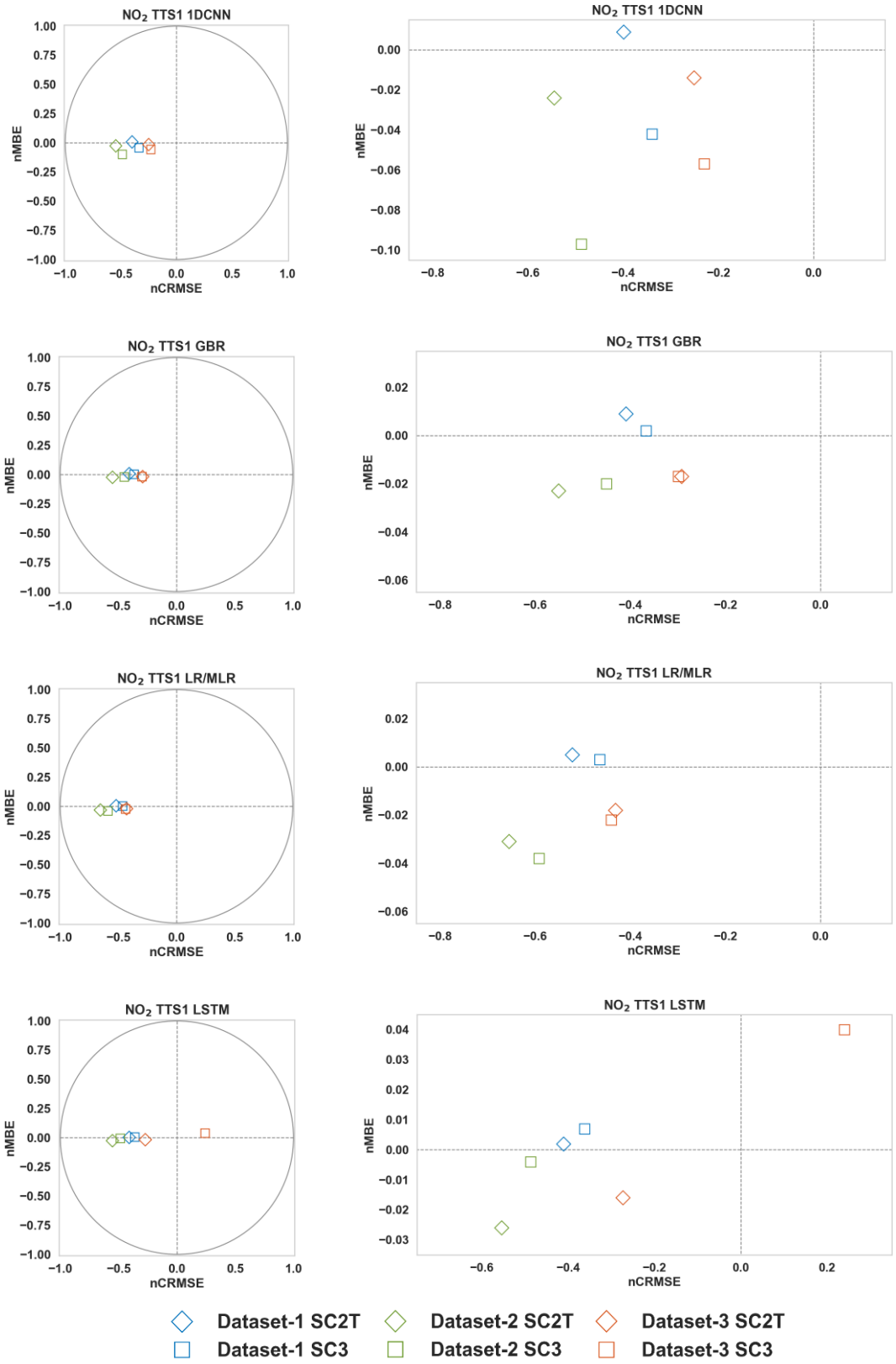
Target diagrams of CO from all three datasets for different algorithms, scenarios: SC3 (same as S1) – raw LCS + Temp + Hum + other gases, SC2T – raw LCS + Temp + Hum + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)

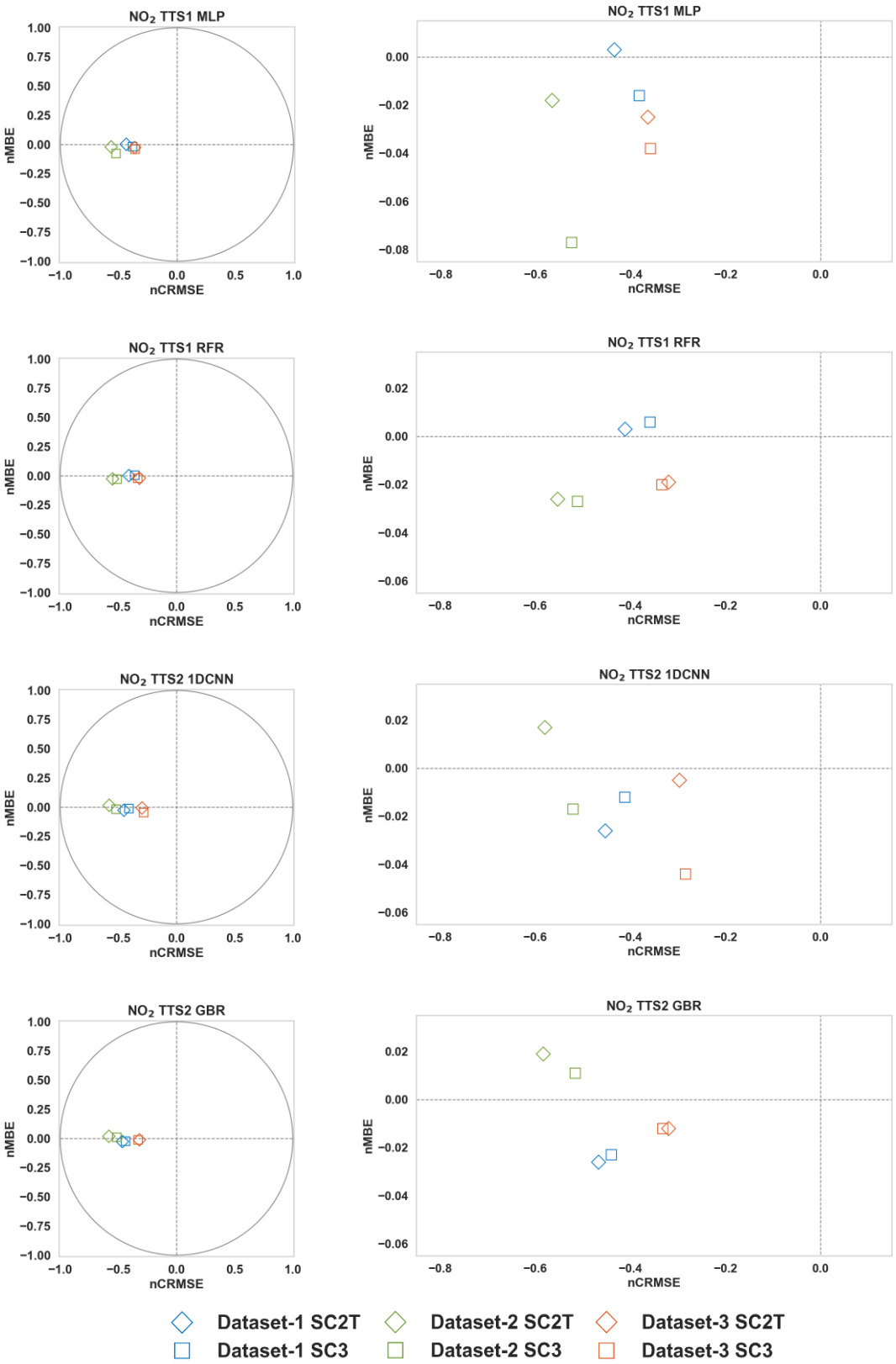


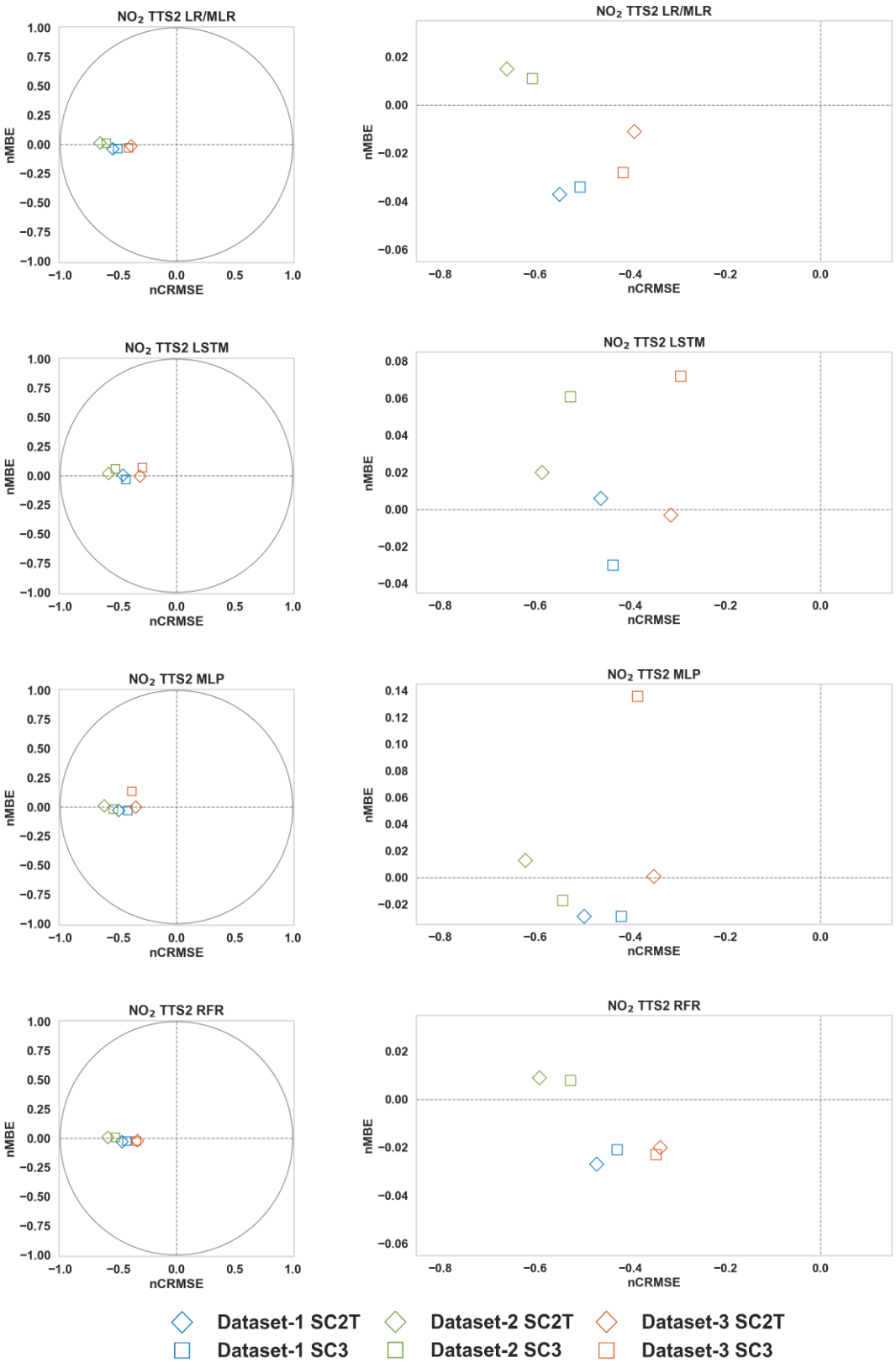

CO TTS1 MLP
CO TTS1 MLP
CO TTS1 RFR
CO TTS1 RFR
CO TTS2 1DCNN
CO TTS2 1DCNN
CO TTS2 GBR
CO TTS2 GBR
nMBE
nCRMSE
Dataset-1 SC2T
Dataset-2 SC2T
Dataset-3 SC2T
Dataset-1 SC3
Dataset-2 SC3
Dataset-3 SC3


CO TTS2 LR/MLR
CO TTS2 LR/MLR
CO TTS2 LSTM
CO TTS2 LSTM
CO TTS2 MLP
CO TTS2 MLP
CO TTS2 RFR
CO TTS2 RFR
nMBE
nCRMSE
Dataset-1 SC2T
Dataset-2 SC2T
Dataset-3 SC2T
Dataset-1 SC3
Dataset-2 SC3
Dataset-3 SC3

Target diagrams of NO_2 from all three datasets for different algorithms, scenarios: SC3 (same as S1) – raw LCS + Temp + Hum + other gases, SC2T – raw LCS + Temp + Hum + N_{days} + Hour, and train test splits (TTS1 – 90:10, TTS2 – 20:80)




NO2 TTS1 MLP
NO2 TTS1 MLP
NO2 TTS1 RFR
NO2 TTS1 RFR
NO2 TTS2 1DCNN
NO2 TTS2 1DCNN
NO2 TTS2 GBR
NO2 TTS2 GBR
nMBE
nCRMSE
Dataset-1 SC2T
Dataset-2 SC2T
Dataset-3 SC2T
Dataset-1 SC3
Dataset-2 SC3
Dataset-3 SC3


NO2 TTS2 LR/MLR
NO2 TTS2 LSTM
NO2 TTS2 MLP
NO2 TTS2 RFR
nMBE
nCRMSE
Dataset-1 SC2T
Dataset-2 SC2T
Dataset-3 SC2T
Dataset-1 SC3
Dataset-2 SC3
Dataset-3 SC3