Master Project HS2020

MSc Environmental Engineering

Performance Analysis of NO, NO_2 Low-cost Sensors and their use in Urban-Scale Air Quality Models

---

Author:
Horim Kim

Supervisor:
Dr. Stephan Henne
Dr. Christoph Hüglin

Zürich, January 2021

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Performance analysis of NO, NO2 low-cost sensors and their use in urban scale air quality models

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Kim	Horim

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Zurich, 31.12.2020	

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.

Abstract

One of the biggest challenges of urban air monitoring is the large spatial and temporal variability of the pollutants. In the present day, the role of low-cost air quality sensors for monitoring is highlighted with regard to their beneficial characteristics; they can be employed in large numbers for higher spatial resolution but can still be cost-effective. This study aims to investigate the performance of the low-cost air quality sensor for nitrogen oxides in urban areas. NO and NO_2 concentrations in the city of Zurich were targeted, and the data were obtained from the aircube sensor developed by *Decentlab GmbH* and calibrated with reference data from the Harkingen monitoring station. Robust linear regression and random forest regression models were utilized for the sensor calibration. The calibration evaluation demonstrated that no significant difference was shown between the two calibration methods concerning the statistical metrics, though the random forest models could not predict the pollutant concentration above a certain limitation. After calibration, the aircube sensors were deployed in four different sites in Zurich. The NO_2 concentration from deployed sensors were compared with passive sampler data, revealing that the sensor performance was degraded in the low pollution regions. An over-estimation of NO_2 concentration in these regions indicated the effect of sensor noise on the prediction. Nonetheless, the comparison with the data from continuous monitoring stations in Zurich claimed that the low-cost sensor performed a good prediction of the pollutant concentration in the urban area. Moreover, relocation of the aircube sensor to the calibration site after deployment, showed that the sensor calibration, as obtained the previous year, could still be utilized for the model prediction. However, some degradation of the performance was clearly shown in the statistic metrics as well as concentration time-series. In conclusion, the performance of the low-cost sensor in this study shows a good potential of the technology to be developed as a sophisticated measure to monitor urban air pollution.

Contents

Abbreviations

CRMSE	Centered Root Mean Square Error
MBE	Mean Bias Error
NABEL	Swiss Federal Air Quality Monitoring Network
nRMSE	normalized Root Mean Square Error
OSTLUFT	Air quality monitoring network for the cantons of eastern Switzerland and the Principality of Liechtenstein
RF	Random Forest regression
RLM	Robust Linear regression
RMSE	Root Mean Square Error

1 Introduction

Significant impacts of urban air pollution on human health are a major concern in many countries. Not only its severe influence on the human respiratory system, but recent studies have also proved that atmospheric pollution adversely affects mental illness, adult diabetes, child growth, and brain development [1-3]. The crisis is especially urgent in major cities of developing countries, for instance, Bangladesh, Pakistan, China, and India [5]. Millions of city dwellers suffer severe exposure to the air contaminants, however, the air quality monitoring is not sufficient to identify exact pollutant sources and flows, furthermore to provide good evidence for environmental policies. One of the biggest challenges of urban air monitoring is the large spatial and temporal variability of the pollutants [6-8]. Even though classic instrumentation is operated to capture the precise air quality data, it cannot be utilized for creating high-resolution information of pollutants. High cost and its requirement of professional operating skills are the main obstacles [6]. Therefore, new techniques for measuring spatially inclusive and comprehensive air quality are crucial for air pollution control in urban areas.

In the present day, the role of low-cost sensors of air pollutants is highlighted with regard to their beneficial characteristics; they can be employed in large numbers for higher spatial resolution but can still be cost-effective. In particular, high data resolution is useful to recognize the pollution hot-spots in urban areas and hence to establish efficient air control policies [7,8]. Low-cost sensors are easy to relocate between two different regions, from an urban area to a rural area, therefore the performance of the sensor can easily be analyzed in various environments [7]. Moreover, the operation is relatively simple without professional skills or additional technical supports [14]. Yet the measurement quality of low-cost sensors has not been proved, thus studies on their reliability and improved calibration strategies are significant for their proper application. The previous research on the air quality low-cost sensors demonstrated their drawbacks. At first, the sensors are significantly affected by meteorologies, such as temperature and relative humidity [6,7,9]. For this reason, field calibration is essential for successful sensor deployment and relocation of the sensor needs additional calibration experiments, and this features can limits the repeatability and reproducibility of low-cost sensors [6]. Secondly, low-cost sensors tend to be cross-sensitive among different ambient pollutants; for example, NO concentration is cross-sensitive to NO_2 concentration and the presence of ozone [9]. At last, the stability of the hardware is still problematic.

This study aims to investigate the performance of NO, NO_2 low-cost sensors in the urban area. For more than a year, low-cost sensors were deployed in the city of Zurich for the research purpose [6,9]. A recent study revealed that Zurich is still exposed to harmful levels of NO_2 [11]. Since NO_2 is an ambient air pollutant that has large spatial-temporal

variation due to its short lifetime, this investigation would be helpful to air quality monitoring in the city [12]. For the low-cost sensor, aircube sensors from *Decentlab GmbH* were utilized in this study. Two calibration methods were applied to get measured data from the deployed sensors; robust linear regression and random forest regression. First, the sensors had been located near the classic air quality monitoring station in Harkingen (6 months) to get data for developing calibration models and then relocated to four distinct areas in Zurich (11 months). Afterward, the performance of low-cost sensors was evaluated by comparing the measured data with the data from precise NO_2 passive samplers. Even though the passive sampler produces data with insufficient temporal resolution (bi-weekly averaged) and it requires additional lab work for the analysis, the observation is known for its accuracy. Therefore, the measurement quality of low-cost sensors can be examined [10]. Furthermore, the calibration model was applied to the raw data obtained from another calibration period (4 months) that the aircubes sensor deployed again in the Harkingen station, in order to additional assessment of the sensor performance and long-term stability. Moreover, continuous NO and NO_2 concentration data from seven monitoring stations around Zurich are compared to explore whether the low-cost sensors accurately capture spatial and temporal inhomogeneity of the pollutants.

2 Materials and Methods

2.1 Data collection and pre-processing

2.1.1 Low-cost sensor data

Aircube sensor data from both calibration and deployment sites were gathered and saved in the aircube database operated by *Decentlab GmbH*. Two NO sensors (NO_00 and NO_01), two NO_2 sensors (NO2_00 and NO2_01), a relative humidity sensor, and a temperature sensor are included in one aircube sensor unit [6]. Detailed sensor description is available on Appendix A.2. With exact date-time information, these six sensor data was contained in the aircube database. Using *R* programming, required data were extracted from the database. It is acknowledged that the low-cost sensor is influenced by its surrounding weather condition, therefore, additional parameters were introduced in the raw sensor data to enhance the sensor measuring accuracy. D_{RH} parameter in equation (1) was introduced in the past research with the same sensor unit and it has been known to mitigate the effects caused by the temporal variation of relative humidity [9].

$$D_{\mathrm{RH}} = \sum_{\Delta t=0}^{-500} \Delta S_{RH}(t + \Delta t) * exp(-\frac{\Delta t}{\Delta t_0}) \tag{1}$$

ΔS_{RH} represents the variation of relative humidity in data acquisition time interval. In this study, four different time steps were chosen for Δt_0 as 60, 90, 120, and 150 minutes. At the last step of pre-processing of the sensor data, the raw data were merged with corresponding reference data for later usage, and any rows contained missing values (*NA* values) in the merged data were deleted for the calibration process.

2.1.2 Measurement sites

Calibration site To develop the calibration model from the reliable concentration data, four aircube sensors (AC009, AC010, AC011, AC012) were installed near Harkingen monitoring site before their deployment. The site Harkingen (HAE: 47.311° N, 7.820° E, 480 m.a.s.l) is a part of the Swiss Federal Air Quality Monitoring Network (NABEL) and is situated near the highway A1, thus the station can be exposed to a wide concentration range of various pollutants [6]. Location map is provided on Appendix A.1.1. In this study, two sensor datasets were obtained from this calibration site; six-month data (29.06.2018 - 12.12.2018; for AC009, the start date was 25.07.2018) were applied to develop the calibration model of low-cost sensors, and four-month data (12.12.2019 - 31.03.2020) were used to verify the model performance in the calibration site. As well as the sensor data, reference data were obtained from the monitoring station for the same periods.

Deployment site The sensors were deployed at four different locations to measure the concentration of air pollutants. (s. Appendix A.1.2) After the six-month installation in the calibration site (29.06.2018 - 12.12.2018), sensors were located in each site for 11 months from 13.12.2018 to 31.10.2019 (For AC010, the measurement started from 14.12.2018). Geographical information with the deployed sensor labels is presented in table 1. It can be highlighted that ZSBS and ZMAN are beside the urban traffic roads in the city of Zurich where the high NO_X concentration presents. On the contrary, ZBLG is the place between the urban residence building and the urban green area, and ZRIS is at the center of the nature area outside of the city. Hence, it was expected that NO_X concentration is lower in this two sites than the others.

Table 1: Four deployment site of low-cost sensors.

Street name (Abbreviation)	Aircube sensor	Coordinate system (LV95)
Im Ris (ZRIS)	AC009	1,246,986 N 2,682,011 E
Bullingerhof 5 (ZBLG)	AC010	1,248,060 N 2,681,593 E
Seebahnstrasse 229 (ZSBS)	AC011	1,248,129 N 2,681,093 E
Manessestrasse 34 (ZMAN)	AC012	1,242,145 N 2,680,665 E

2.1.3 Reference data

The reference data for NO and NO_2 concentration was measured from the Harkingen monitoring station. The measurement was based on the intensity of chemiluminescence radiation from the reaction of NO with ozone, a detailed information can be found in the technical report from NABEL [13]. Obtained from the NABEL database, the collected data in three conditions were cleaned; 1) data obtained when the calibration gas was injected to the monitoring instrumentation, 2) when no gases were measured, 3) data in the periods of maintenance. Duplication of the data and missing data (*NA* values) was also checked and neglected. Moreover, the original date-time in Central European Time (CET) was converted to Coordinated Universal Time (UTC). After the refinement, 1-minute data were aggregated for the 10 minutes time resolution to correspond to the aircube sensor data. At last, NO and NO_2 concentration information in the final version were merged into the aircube sensor data for the calibration process.

2.1.4 NO_2 passive sampler data

NO_2 passive diffusion samplers were located in four deployment sites as aircube sensors (ZRIS, ZBLG, ZSBS, ZMAN). The sampler network was operated by the department of Environment and Health Protection of the City of Zurich (UGZ) [10]. Averaged values of NO_2 concentration was provided bi-weekly from 04.12.2018 to 05.11.2019; this means

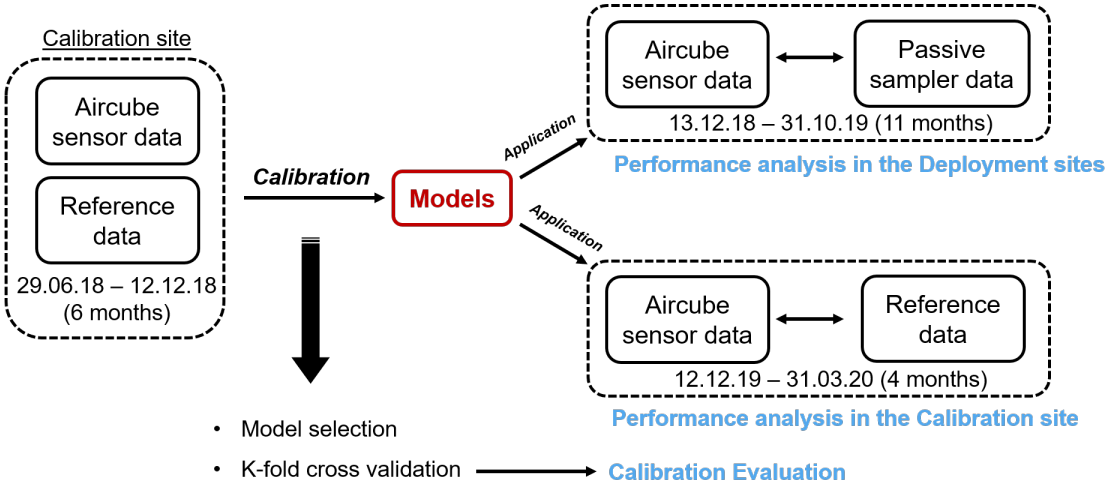
that 24 values of concentration were considered in each site. Given concentrations have a unit of $\mu g\,m^{-3}$, hence the unit was converted to *ppb* with a factor of 1.9125.

2.1.5 Continuous monitoring data

Continuous monitoring data of NO and NO_2 concentration were provided by NABEL and OSTLUFT. Detailed information about the monitoring stations and the sensors are found in the appendix A.1.3. The data from seven monitoring stations (DUE, KAS, SCH, HEU, STB, RGS, and OPK) are available from November 2018 (January 2018 for SCH) until the end of 2019, and they have different time resolution from 10 minutes to 60 minutes. In all stations, the measurement was based on the intensity of chemiluminescence radiation as the Harkingen monitoring station [13]. The concentration from OSTLUFT had a unit of $\mu g\,m^{-3}$, hence it was converted to ppb.

The given monitoring data were compared with the aircube sensor concentration from four deployment sites. The objective is to examine the aircube sensor performance in the deployment sites. Histogram for the NO and NO_2 concentration of each site were extracted from the time-series data, and it was utilized to evaluate the pollution in the locations.

2.2 Calibration process



2.2.1 Calibration methods

Two calibration methods were utilized and evaluated in this study: Robust linear regression (RLM) and Random forest regression (RF). Raw data of aircube sensors along with comparable reference data are applied to these methods to estimate the NO and NO_2

concentrations. In principle, seven electrochemical signals from sensor raw data were considered as parameters; V_{Sensor}, $V_{Tempertaure}$, $V_{Relative\ humidity}$, and four $V_{D_{RH}(}\Delta\ t_0$ as 60, 90, 120, and 150 minutes). These values comprise the formula of the calibration models. V_{Sensor} represents the signal of two NO sensors and two NO_2 sensors. The concentration sensors in an aircube sensor unit were calibrated independently.

Owing to the multifarious combination of these parameters to a model formula, 26 RLM models and 9 RF models were evaluated primarily. (All of the model formulas are introduced in Appendix B.) Among the 35 models, three RLM models and the two RF models were selected for further investigation on the study. The model selection process was executed based on the target diagram of the model residuals. Detailed information about this evaluation approach is demonstrated in section 2.2.4.

Concerning the process of the model selection, 80% of the sensor data in six-month (29.06.2018 - 12.12.2018) was randomly chosen and used for the model training and the other 20% was applied to model testing to obtain the information for the target diagram. There are 16 concentration sensors in overall; four aircube sensor units with four concentration sensors respectively. For a selection criteria, the 35 models were ranked based on the distance of the target diagram in each sensor; the least distance in the target diagram ranked the highest. Furthermore, these rankings from each sensor (16 ranking tables) were then combined into one rating table. As a results, the final five models for the study were determined according to the final ranking.

Robust linear regression Multiple linear regression is one of the most important methods in the field of calibration. Indeed, the effectiveness of this methodology in the air quality low-cost sensor has been already confirmed by several studies [6,7,18,19]. Robust regression is a technique that reduces the model distortion induced from unusual observations [19]. By limiting outliers to get involved in the modeling, the 'robustness' prevents their extreme impacts.

The `rlm()` function from *R* package 'MASS' was used for the calibration [20]. The function fits a linear model by robust regression using an M-estimator. All the parameters in the function were brought from their default values [20]. Huber's loss function with k value of 1.345 was taken for the psi function, and *maxit* value was set as 1,000. The *maxit* value indicates the maximum limit on the number of iterations [20].

Random forest regression A random forest is a machine learning method in which numerous decision trees are examined to identify a superior model of a classification or regression. Each node in a decision tree is split by using the best option among a subset of predictors that are randomly chosen at that node [21]. Subsequently, multiple decision trees are ranked and the best option was selected as an output. In particular,

a random forest regression method has been utilized as a low-cost sensor calibration in several studies [6-8]. RF models shows great performance in these studies, however, it was shown in previous research that the overfitting of the model may occur during calibration [7]. Moreover, the method could not predict values which are beyond the range of model training. [8].

In the present work, `randomForest()` function in *R* package 'randomForest' was employed for the calibration. All the other parameters were taken as default values of the R function [22], however, two major parameters for the formation of decision trees were strategically determined. First of all, a *ntree* value, a parameter that decides the number of decision trees, was chosen as 1,000. The choice relies on a literature [23], as well as a limit on the computational power. Secondly, a *nodesize* value was decided by the numerical assessment of the model performance. With 9 different node sizes (*nodesize* as 5, 10, 20, 30, 40, 50, 100, 200, and 300) in each function, nine models were evaluated based on the target diagram. As a consequence, the node size of 100 was selected because the model of this parameter had the least normalized RMSE value. The detailed results of the node size parameter decision are presented in Section 3.1.1.

2.2.2 K-fold cross-validation

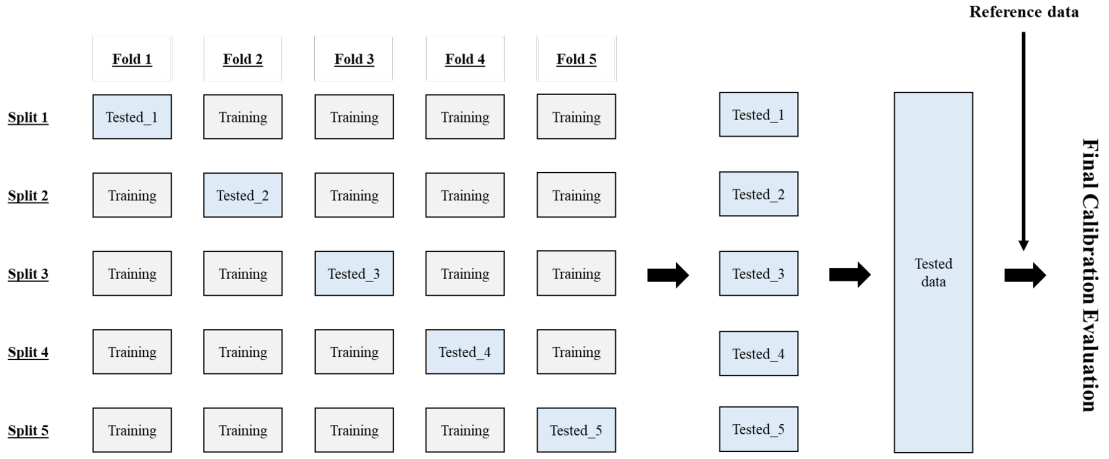


Figure 1: Scheme of k-fold cross-validation process

Five chosen calibration models from the selection process were evaluated by K-fold cross-validation. The method is widely used to estimate the prediction error and model accuracy. The size of the tested portion increases, hence the calibration evaluation can be more precise [24]. In the present study, the number of the fold (K) is 5 as a typical value, due to the computational limitation [25]. Pre-processed low-cost sensor data were randomly split into the five different sub-groups, and four sub-groups of data (80% of total) were used as training data in each fold, and the rest one group tested the model. As a result

of 5-fold cross-validation, five predicted concentration data were combined again into one dataset. (s. Figure 1) The dataset were then evaluated with corresponding reference NO and NO_2 concentrations.

2.2.3 Model application to sensor data

Calibration models were established by training the model with the data from the six months in the calibration site. Developed models were then applied to two target sensor data, which are the deployed sensor data (13.12.2018 to 31.10.2019) and the sensor data in another calibration site period (12.12.2019 - 31.03.2020). An objective of the first implementation is to analyze the aircube sensor performance in various locations in the city with different pollutant concentrations. On the other hand, the second application has the purpose of the investigation on a relocation of the low-cost sensors in the calibration site. In this application, reference data in 1-minute resolution were obtained from the Harkingen monitoring station so that it allowed reliable evaluation of the model performance.

2.2.4 Evaluation approach

Two approaches were considered in the calibration evaluation: (1) Target diagram visualizing various statistic measures, and (2) Coefficient of determination that examines the performance of the regression models. Both methods are calculated by using `tdStats()` function in *R* package 'tdr' [26].

Target diagram The methodology has been utilized in the research on air quality low-cost sensors [6,7,18]. The basic structure of the target diagram (see figure 2) is straightforward; On the x-axis, the centered root means square error (CRMSE) normalized by the standard deviation of reference data ($\sigma_{\text{reference}}$) is stated, whereas the normalized mean bias error (MBE) is stated on the y-axis. The statistical definition and the formula of measures are elaborated in Appendix C. A significant feature of this diagram is that the figure contains multiple information about the characteristics of models and their residuals. Below are the four diagnostic measures that can obtain from the plot [7]. In particular, the

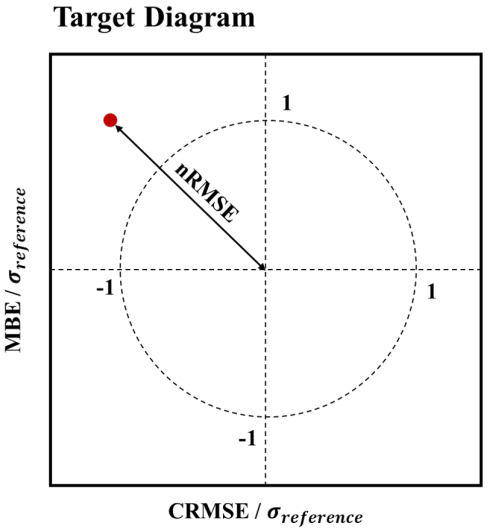


Figure 2: A structure of the target diagram. The red point indicates an example of a value position in the diagram.

nRMSE value from the target diagram was explored in the model selection and calibration evaluation processes.

(a) A vector distance between the coordinate and the origin represents the normalized RMSE (nRMSE, RMSE/$\sigma_{\text{reference}}$) value.

(b) The model prediction is overestimated (MBE > 0) or underestimated (MBE < 0).

(c) The standard deviation of the model prediction is larger (CRMSE/$\sigma_{\text{reference}} > 0$) or smaller (CRMSE/$\sigma_{\text{reference}} < 0$) than that of the reference data.

(d) The variance value of the model residuals is larger (outside of the circle of radius 1) or smaller (inside of the circle of radius 1) than that of the reference measurements.

Coefficient of determination ($\mathbf{R^2}$) In general, R^2 quantifies the ability of the model prediction in the linear regression model [27]. Higher value of R^2 means a good-fit of a model data to the observation data. In this study, this value was used to determine how well the calibrated sensor data are corresponding to the reference measurements and the modeling outputs.

3 Results and Discussion

3.1 Sensor calibration

3.1.1 Parameter decision : Node size in a random forest model

The node size parameter in the random forest model indicates the minimum number of observations in a terminal node [23]. A low node size leads to the large depth of trees, thus the more branches created and the computation time for regression increases [23]. Therefore, an appropriate node size should be tuned before implementing the models to get optimal results with short time consumption.

In this study, the node size parameter was determined by an experiment to identify the value which has the least nRMSE in the target diagram of various models. For the experiment, four sensors from each aircube unit were considered: NO_00 sensor in AC009, NO_01 sensor in AC010, NO2_00 sensor in AC011, NO2_01 sensor in AC012. The *ntree* parameter in the regression function was fixed as 100 instead of 1,000 due to the limited computational time. As a result, nRMSE values from the target diagram were obtained and the results are presented in the appendix D.1. Node size with the least nRMSE value in each model in four sensors is found in Table 2.

Table 2: The result of the experiment for the node size paramater decision. For each sensor and model, the parameter values with the least nRMSE in the target diagram are stated in the table.

Unit	**Sensor**	RF_1	RF_2	RF_3	RF_4	RF_5	RF_6	RF_7	RF_8	RF_9
AC009	NO_00	300	100	40	30	20	30	100	20	200
AC010	NO_01	300	100	40	30	30	30	100	50	100
AC011	NO2_00	300	100	20	30	20	30	100	10	100
AC012	NO2_01	300	100	30	30	30	20	100	20	100

The parameter value was selected as 100 for further investigation since it is one of the most frequent values as the least nRMSE in table 2. Indeed, the value of 30 was as frequent as 100 in the table, 100 was preferred since it requires short computation time in the calibration. However, the determination was not substantial to the regression model for two reasons. Firstly, the nRMSE was not sensitive to the node size in random forest regression as shown in the appendix D.1. In most experiments, the variation of the nRMSE value is less than 0.05. Secondly, the least nRMSE is not identical in all models, but is dependent on the model complexity and included parameters. When the model considered both temperature and relative humidity as parameters, a lower node size (less than 100) was preferred. Additional investigation is required for the parameter of random forest regression yet it is out of scope in this project.

3.1.2 Results of model selection

Simulations with 35 pre-defined calibration models were conducted in each sensor of four aircube units. The nRMSE values as a result of the simulation evaluation in each model in all sensors are elaborated in the appendix D.2. The selection was determined by the ranking of the least 'averaged nRMSE' among the models. 16 nRMSE values from each sensor were averaged to represent the models, and three RLM models and two RF models with lower 'averaged nRMSE' were selected. If two chosen models have the same formula but only the Δt_0 of D_{RH} is different, the higher-ranked model was taken and the other was substituted by other models. Selected models with new labels are found in table 3.

Table 3: Five selected calibration models with new labels from the model selection.

Old labels	New labels	Model formula
RLM_26	RLM_A	Reference = function(V_{sensor}, T, V_{sensor}*T, V_{sensor}*T^2, D_{RH_150})
RLM_17	RLM_B	Reference = function(V_{sensor}, T, RH, D_{RH_150})
RLM_22	RLM_C	Reference = function(V_{sensor}, T, V_{sensor}*T, D_{RH_150})
RF_6	RF_A	Reference = function(V_{sensor}, T, D_{RH_150})
RF_8	RF_B	Reference = function(V_{sensor}, T, RH)

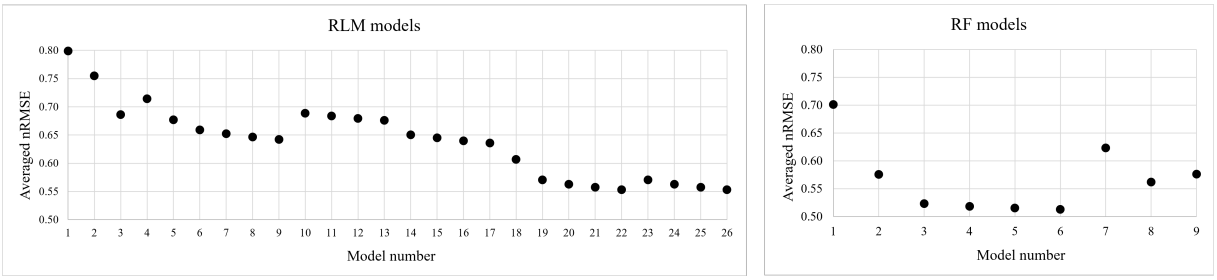


Figure 3: The graphs of averaged nRMSE values in each calibration model. Averaging was implemented to each calibration models by concerning the nRMSE values from all sensors.

Overall results of the model selection process describe the feature of the model parameters as well as the feature of the aircube sensor. Firstly, it is prominent that the models which ignore the parameter of temperature signal (T) have higher nRMSE values than the models with the parameter (s. RLM #1, #2, #4, and #10 - #13 models and RF #1 and #7 models in figure 3). This is due to the temperature dependence of the aircube sensor; the temperature varied from -4°C to 35°C in the given period, and the temperature dependence of the sensor is elaborated in the performance data sheet from *Decentlab* [29]. Secondly, the presence of D_{RH} parameter in the model diminishes the nRMSE values more than the relative humidity (RH) parameter (s. RLM #5 - #9 models and RF #3 - #6 and #8 models in figure 3). The effect of relative humidity on the aircube gas sensor was

proved in previous research [28], however, this result can supplement that the temporal variation of relative humidity is more influential than the present relative humidity itself. Furthermore, it has been explored that the model with D_{RH} parameter of higher Δt_0 has the lower nRMSE. Nevertheless, this trend is not eminent in all possible Δt_0. The previous research [28] experimented the simulation with more various Δt_0 (40, 50, 60, 80, 100, 120, 140, 160, and 180), and the RMSE was the lowest at Δt_0 of 120 minutes. Lastly, the existence of the parameter $V_{sensor} * T$ in RLM models reduces the nRMSE value. Comparing RLM #6 - #9 models with RLM #19 - #22 models, the parameter reduces ∼0.1 nRMSE. The parameter may mitigate the temperature sensitivity of the sensor, but additional study is required for understanding its effect.

3.1.3 Calibration evaluation of K-fold cross validation

To assess the performance of five chosen models from the selection, two metrics were calculated for each aircube sensors. The metrics are nRMSE from the target diagram, and R^2 from the correlation plot between calibrated and reference concentrations. Table 4 summarised the metric values by the calibration model types and by the target pollutants.

Table 4: Evaluation metrics for the sensor calibration. The table contained averaged values of metrics by two criteria. All the values of nRMSE and R^2, target diagrams, and the correlation plots for each sensors are presented in the appendix D.3.

	Averaged by calibration model types					Averaged by target pollutants	
	RLM_A	RLM_B	RLM_C	RF_A	RF_B	NO	NO_2
nRMSE	0.55	0.62	0.56	0.63	0.57	0.48	0.69
R^2	0.71	0.63	0.70	0.74	0.68	0.77	0.61

The result demonstrates that no significant difference between RLM and RF models is shown in terms of nRMSE and R^2 values. The differences are less than 0.1 in both parameters. In addition, target diagrams in the appendix D.3.1 indicate that MBE is less than 0.01 in all sensors, though all the CRMSE/$\sigma_{\text{reference}}$ values are negative, which means that the standard deviation of the model prediction is smaller than that of the reference data. However, the low standard deviation in the predicted concentration is unsurprising, because the prediction could not completely estimate the extreme values of the reference data.

Nevertheless, the averaged parameter by target pollutants shows a considerable difference. The averaged value of NO sensors has lower nRMSE and higher R^2, and this tendency is not only shown in the averaged representatives but also in the metrics of respective sensors (s. table 18 and 19 in appendix D.3). A reason that the sensor performance is

worse in NO_2 than NO would be caused by the physical feature of the sensor. The possible noise range of the NO and NO_2 sensor in the aircube sensor is $\pm$ 15 ppb (s. table 8 in the appendix A.2), and the reference concentration of NO_2 is mostly in the range that is largely affected by the sensor noise as shown in figure 4. To be specific, 80.7% of the NO_2 concentration value in the given periods are below 30 ppb. Not only to the NO_2 sensors, the noise in low reference concentration would affect the performance degradation of all sensors in the aircube unit. A comparable outcome can be seen on the sensor performance in the deployment sites in Section 3.2.1.

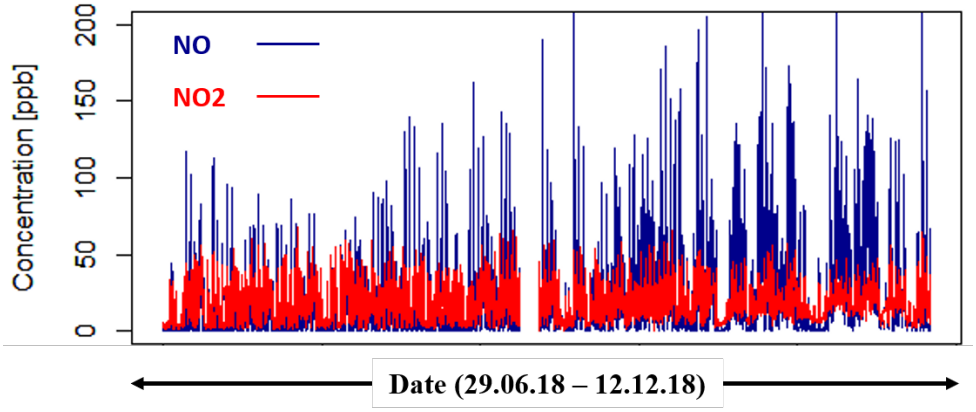
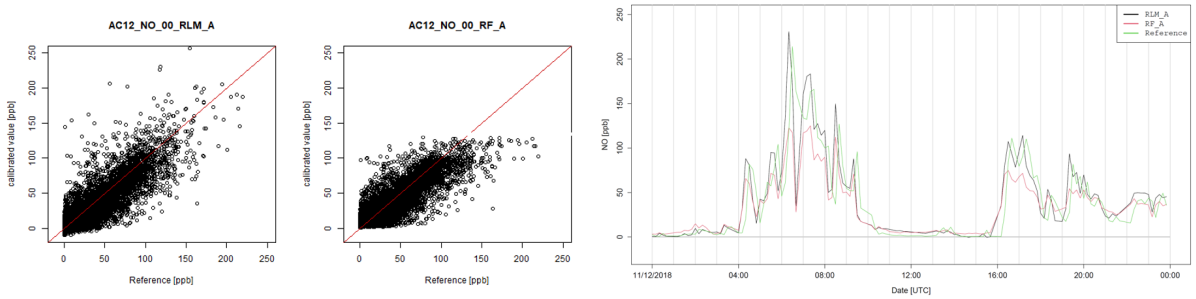
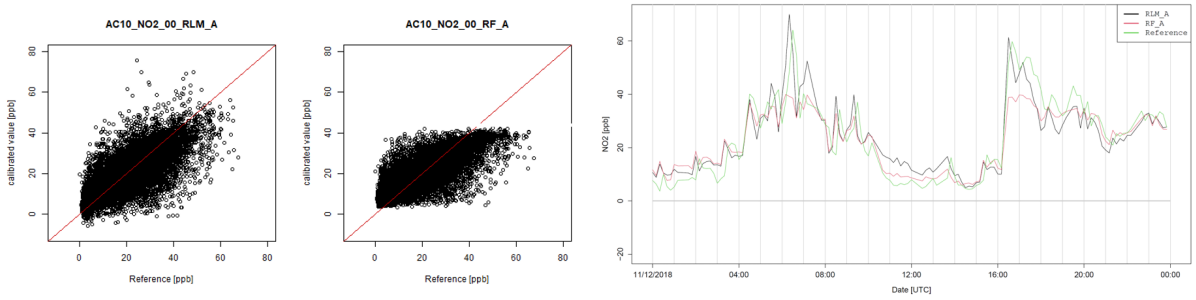


Figure 4: Reference NO and NO_2 concentration from the Harkingen monitoring station in the six-month calibration period (29.06.18 - 12.12.18).

Furthermore, the scatter plot between the sensor concentration and the corresponding reference concentration depicts that the prediction from RF models have upper limitation in all sensors (s. figure 18, 19, 20, and 21 in appendix D.3.2). Two examples are presented in figure 5. The scatter plot of RF_A model of the NO sensor illustrates the upper limits of NO concentration as ~130 ppb, whereas that of NO_2 sensor demonstrates the upper limits of NO_2 as ~40 ppb. These limitations did not appear in any RLM models, but only in RF models. Moreover, the concentration time-series beside the scatter plots show that the RF_A model could not represent the concentration peak, and this type of misprediction is also observed in other calibration periods. The concentration time-series of full-time duration are found on the supplementary documents of this project. As a result, the upper limit in the prediction is a deficiency of the RF regression method.



(a) NO_00 sensor of AC012



(b) NO2_00 sensor of AC010

Figure 5: The scatter plots of RLM_A and RF_A model prediction in NO_00 sensor of AC012 and NO2_00 sensor of AC010, with their daily concentration time-series in 11.12.2018. The date was chosen because the concentration peak exist and the upper limitation of the RF model prediction is displayed.

3.2 Performance analysis of the low-cost sensor

3.2.1 Comparison to passive sampler data in the deployment sites

Four passive samplers located in the deployment sites provided 24 values of the bi-weekly averaged concentration of NO_2, and the data was compared with the concentration from corresponding low-cost sensors. Scatter plots obtained from each sensor are presented in figure 6. In general, an over-estimation of the concentration in the aircube sensor is shown in both AC009 and AC010, whereas the under-estimation is shown in the AC012. AC011 performs the best among four deployed sensors based the nRMSE value in the table 20 in appendix E.1.

The demonstration from the plots argues that the sensor performance reflects the characteristics of its location. First, AC009 and AC010 were placed in the green area (s. the picture of the deployment sites in figure 14 in appendix A.1.2), thus they were exposed to the relatively lower concentration of NO_2 contrary to the other locations. The passive sampler concentration was less than 10 ppb in ZRIS(AC009) and was around 5 to 20 ppb in ZBLG(AC010). Again, the sensor performance would be influenced largely by the sensor noise when the reference concentration is low, as observed in the section 3.1.3. The weekly time-series of the aircube sensor were depicted in figure 7. In figure7(a), the dissimilarity among the five models was relatively small during the concentration peak,

however, it increased during the low concentration period in 23.02.19 - 24.02.19. No significant distinction between the RLM and RF model appears in the time-series, though the bi-weekly averaged concentration of RLM models is more closed to the passive sampler data in the scatter plots.

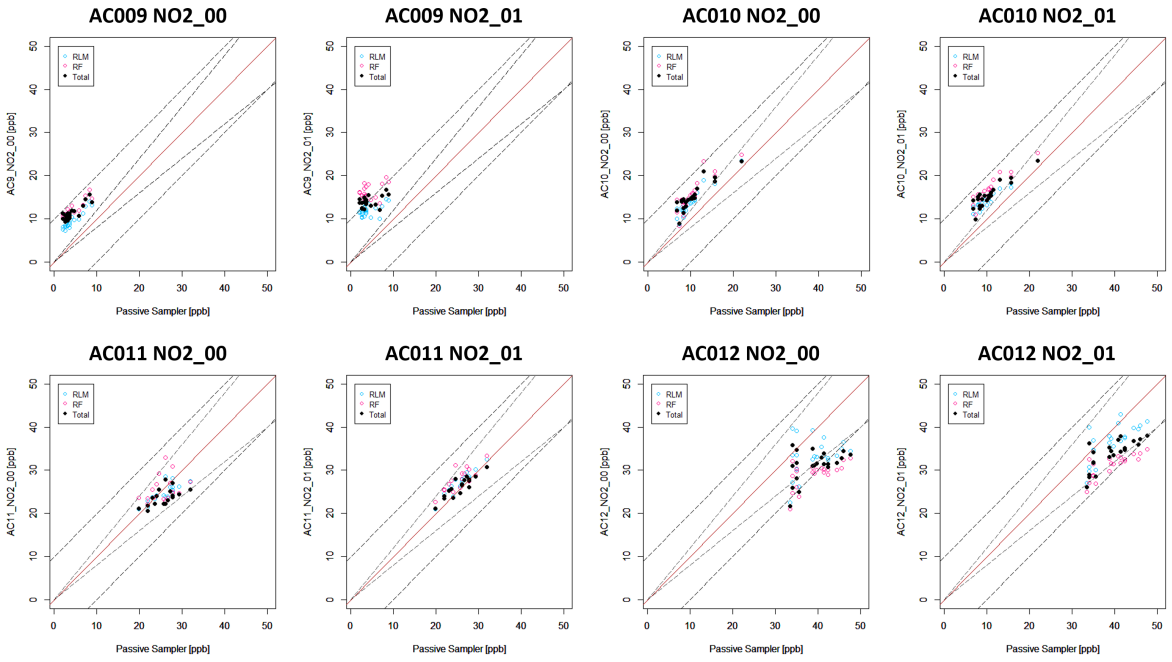
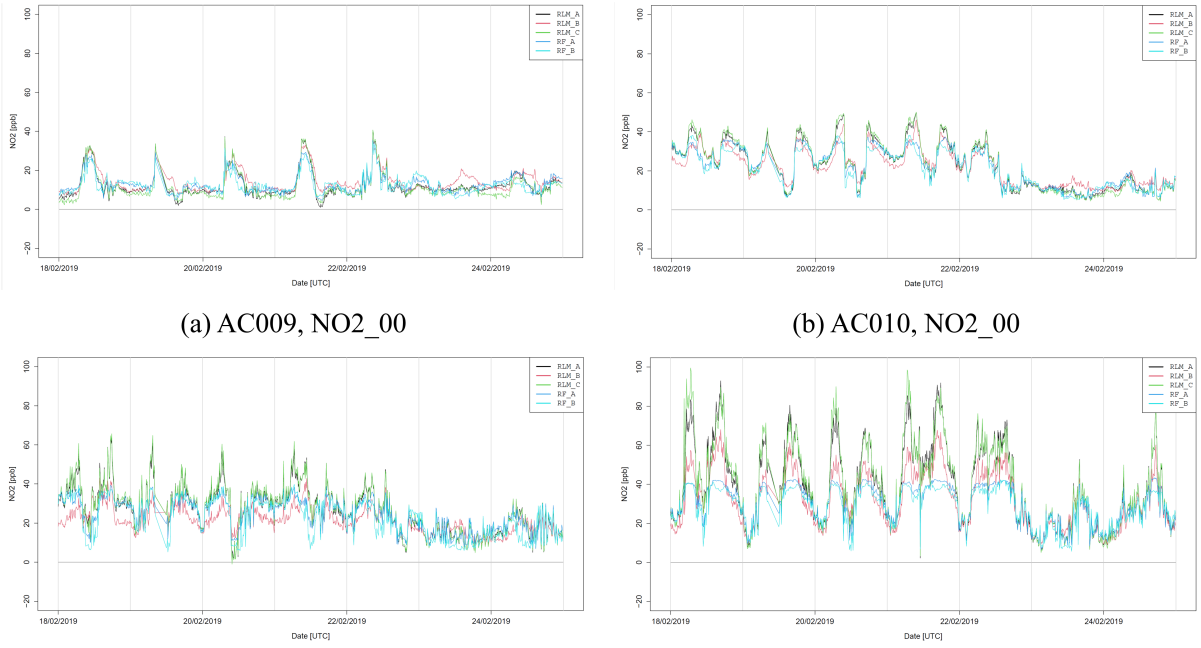


Figure 6: Scatter plots between the NO_2 concentration from the passive sampler and that from the aircube sensors. Two plots were generated from two NO_2 sensors in each sensor unit.

On the contrary to AC009 and AC010, the other two sensors were located beside the urban traffic roads in the city of Zurich (s. the picture of the deployment sites in figure 14 in appendix A.1.2). Hence AC011 and AC012 were exposed to a high enough concentration leads to good predictions. The peak concentration of AC011 is ∼80 ppb and that of AC012 is ∼100 ppb according to their weekly concentration time-series. AC011 shows good performance in both RLM and RF models and no difference was identified between the models. Nevertheless, the RF models in AC012 show considerable underestimation compared to the passive sampler concentration, and this can be also revealed in the time-series in figure 7(d). The problems have already been detected in the calibration evaluation part; the NO_2 concentration in the RF models could not represent high pollutant concentration over certain points. The upper limitation is ∼ 40 ppb in the figure 7(d), which is identical to the corresponding value in the calibration evaluation. A clear explanation of this drawback in the performance of RF models was not fully discovered in this study. However, it can be assumed that a lack of model training in high pollutant concentration caused the maximum limits in the RF model prediction. A disadvantage of the RF regression was explained that the model could not estimate the value beyond

the concentration range of model training data. Nevertheless, additional investigation is required to confirm this assumption.



(a) AC009, NO2_00

(b) AC010, NO2_00

(c) AC011, NO2_00

(d) AC012, NO2_00

Figure 7: Weekly time-series of aircube sensor concentration in the deployment sites. A week (18.02.19 - 24.02.19) was illustrated as an example.

3.2.2 Comparison to reference data in the calibration site

After the deployment in four sites, aircube units were again installed in the calibration site for 4 months (12.12.19 - 31.03.20). The relocation has an objective to describe the model performance one year after the calibration, thus the sensor performance in this period was compared with the calibration evaluation. Averaged evaluation metrics are shown in the table 5. Comparing the nRMSE with that of the table 4, the difference is not substantial. In general, nRMSE value increases only 9.5 % from the previous period. Target diagram in the appendix E.2.1 also stated the minor performance degradation. The sensors have small MBE and only some of the models have positive CRMSE/$\sigma_{\text{reference}}$ on contrary to the diagram from the calibration process.

Table 5: Evaluation metrics for the sensor performance in the calibration site. The table contained averaged values of metrics by two criteria. All the values of nRMSE and R^2, target diagrams, and the correlation plots for each sensors are presented in the appendix E.2

	Averaged by calibration model types					Averaged by target pollutants	
	RLM_A	RLM_B	RLM_C	RF_A	RF_B	NO	NO_2
nRMSE	0.64	0.67	0.63	0.68	0.59	0.49	0.80
R^2	0.64	0.55	0.64	0.53	0.65	0.79	0.41

Nevertheless, the R^2 values and the correlation plots in E.2.2 demonstrated the unsatisfactory performance of the sensor, and it is especially distinctive in NO_2 concentration. Examples of the plots are shown in figure 8. The figure shows that severe under-estimation was derived in the sensor, and this trend is also found in other aircube units. A case of the RLM_A and RF_A model of this sensor were investigated in detail and the result was demonstrated in figure 9. The figure identified that the under-estimation significantly occurred from mid-February to mid-March, but no evident reason was detected from the time-series. Times-series of other sensors can be found in the digital supplementary material. Further analysis would be required to identify the rationale behind this problem.

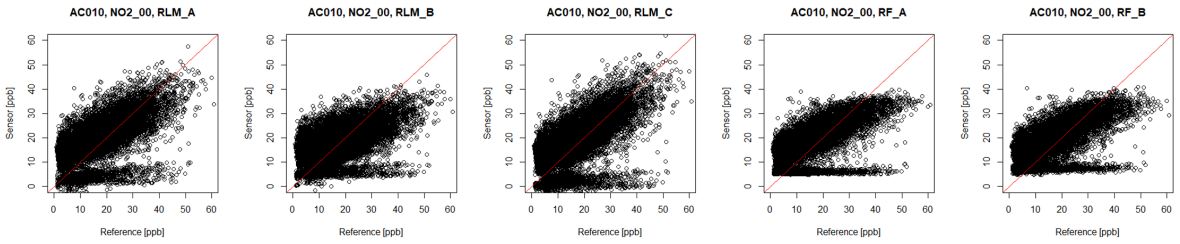
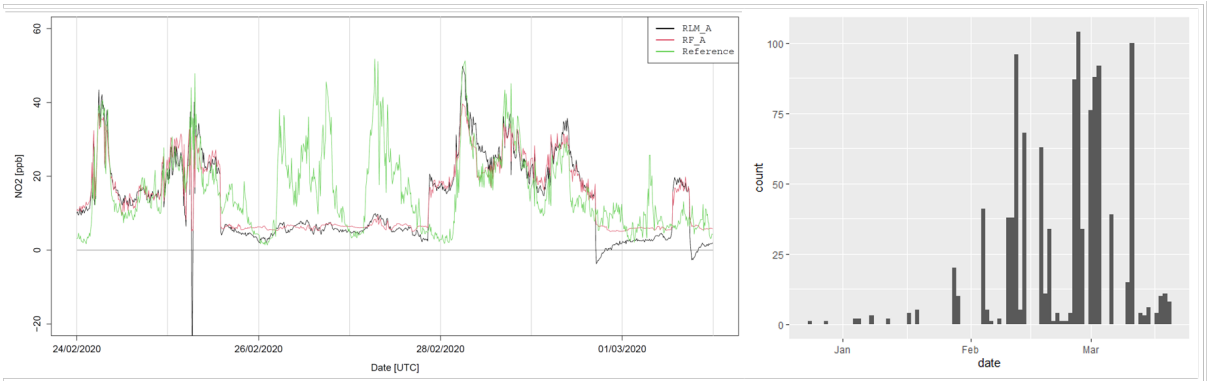


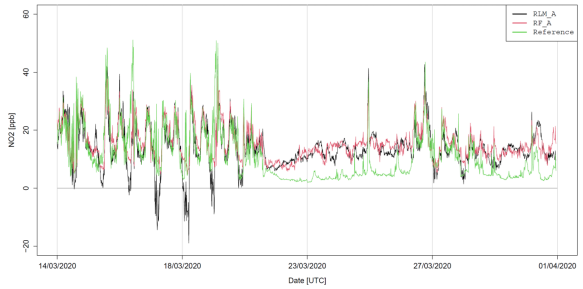
Figure 8: Scatter plots between sensor and reference concentration of NO2_00, AC010.



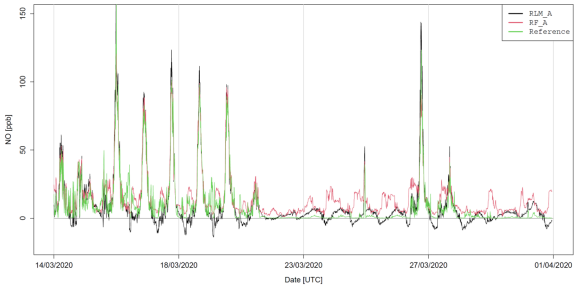
(a) NO2 concentration time-series (b) Number of days in under-estimation 'zone'

Figure 9: Example of NO_2 concentration time-series. The data from RLM_A and RF_A model with the reference data of NO2_00, AC010 are illustrated. Figure (b) shows the number of day in each periods that the sensor concentration are located in the severe under-estimation 'zone' in the scatter plot of RF_A models in figure 8. The 'zone' here indicates the long tails in the scatter plot where the sensor cocentration is less than 10 ppb and the reference concentration is bigger than 10 ppb.

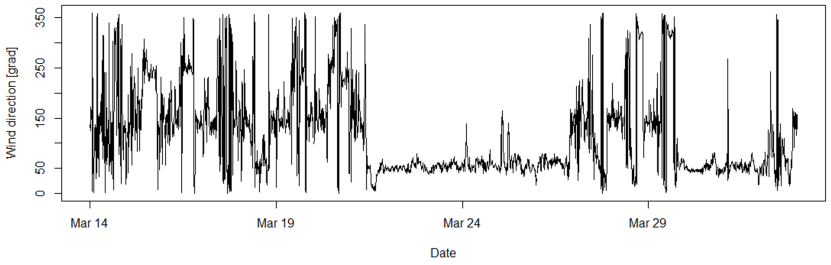
In addition to the under-estimation, an over-estimation was also substantial in the model prediction. In particular, the over-prediction happened when the reference NO and NO_2 concentration is less than 10 ppb. This type of the bad performance of the sensor has already been discussed in section 3.1.3. and 3.2.1., yet the new case is illustrated in figure 10(a) and 10(b). Indeed, a substantial part of the low reference concentration was caused by the wind direction in the calibration site. As shown in the wind direction time series in figure 10(c), the over-prediction with low reference concentration is presented when the wind direction was between 35-80 degrees. This corresponds to wind from the north-east direction, which means advection from an agricultural area with very little NOx emissions instead of direct advection from the highway with dense traffic from the opposite direction (s. the picture of the calibration site in figure 12 in appendix A.1.1). Thus, it is concluded that the meteorological condition of the calibration site affected the sensor performance.



(a) NO2 concentration time-series



(b) NO concentration time-series



(c) Wind-direction in the calibration site

Figure 10: Concentration time-series of RLM_A and RF_A model with corresponding reference data. The sensor is NO2_00, AC012. Time period of the plot is 14.03.20 - 01.04.20. Figure (c) represent the wind-direction of the monitoring station in the calibration site during the same period.

Overall, the model application to the sensor after a year confirmed that the sensor performance is still in the acceptable range regarding the statistical metrics in the target diagram. However, unlike the calibration evaluation, the scatter plots with the reference revealed the degradation of the performance.

3.2.3 Comparison to continuous monitoring data in Zurich

The histogram of NO and NO_2 concentration from four deployment sites and seven monitoring stations are presented in figure 27, 28, and 29 in appendix E.3. In addition, the mean concentrations with statistic metrics of two pollutants are presented in table 6. In general, every histogram depicts an analogous trend. The figures demonstrated that the NO concentration is highly positive-skewed, whereas the NO_2 concentration has a smaller coefficient of skewness. Moreover, the coefficient of skewness increases, and the standard deviation decreases when the location has a low pollutant concentration. This is observed by comparing less polluted deployment sites(AC009 and AC010) with the others(AC011 and AC012).

Based on the mean concentration, the histogram of aircube sensor data can be matched to that of monitoring station data. In other words, AC009 can be coupled with DUE or KAS, AC010 with STB, and AC011 with RGS. In these groups, the coefficient of skewness is also similar to each other, while the aircube sensor data has a lower standard deviation. The

latter was due to the lack of the ability of the calibration model to predict the temporal variation (s. Section 3.1.3. and 3.2.2.) The mean concentration of NO_2 from the passive samplers are also presented in the table, and the under-estimation in AC009 and AC010 is also captured.

Overall, the comparison with the continuous monitoring data suggested that the aircube sensor can detect the comparable concentration of pollutants as the reference stations non only in the city center area, but also in less-polluted urban green areas like AC009 and AC010. Yet, it is discovered that no comparable continuous monitoring station was as polluted as AC012 (ZMAN). Indeed, NO concentration in ZMAN is 109.7 % higher than that of the highest polluted station (RGS), and 44.0 % for NO_2 concentration. Hence, this analysis insists that an additional air quality monitoring system is required to detect the pollution hot-spots in Zurich.

Table 6: Mean concentration, the moment coefficient of Skewness, and the standard deviation from the NO and NO_2 data of deployment sties and the continuous monitoring stations in Zurich. In addition, the mean concentration of NO_2 from the passive samplers in the deployment sites are presented as PS009 to PS 012

	Mean concentration [ppb]		Moment coefficient of skewness		Standard deviation	
	NO	NO2	NO	NO2	NO	NO2
AC009 (ZRIS)	4.94	14.00	5.67	1.76	4.17	3.94
AC010 (ZBLG)	7.16	16.15	4.75	0.91	8.82	6.91
AC011 (ZSBS)	25.61	25.87	2.04	0.54	28.11	8.33
AC012 (ZMAN)	53.18	33.94	1.30	-0.10	44.59	12.05
DUE	4.37	11.39	4.43	1.21	10.62	9.09
HEU	1.47	6.77	7.87	2.12	3.68	5.70
KAS	3.82	12.87	5.73	1.23	9.27	9.20
OPK	15.96	18.00	2.08	0.89	17.91	10.71
RGS	25.36	23.57	2.06	0.66	25.09	10.75
SCH	17.07	18.86	2.72	0.64	21.03	10.16
STB	8.01	13.78	7.74	2.13	15.23	9.96
PS009	-	4.16	-	-	-	-
PS010	-	10.59	-	-	-	-
PS011	-	25.05	-	-	-	-
PS012	-	39.24	-	-	-	-

4 Conclusion

The current study investigated the calibration of four low-cost air quality sensors and examined the sensor performance in multiple locations by comparison with reference observations. In the sensor calibration, two regression methods, robust linear regression and random forest regression, were utilized. A model parameter and the model formula were tested to identify the best calibration model, and three RLM models and two RF models were selected for further performance analysis. The model selection revealed that the low-cost sensor was vulnerable to environmental conditions. The models that contained parameters of temperature and relative humidity performed better than the models without them. Moreover, the current simulation results suggested that D_{RH} parameter with Δt_0 value of 120 accomplished the lowest nRMSE, however, additional studies are required to identify its influence on the calibration process.

The calibration evaluation demonstrated that no significant difference was shown between RLM and RF models concerning the statistical metrics, though the RF models could not predict the pollutant concentration above a certain limitation. The limitation was ∼130 ppb for NO and ∼40 ppb for NO_2, and this boundary was also shown in the model performance in the deployment sites. It was presumed that the limitation was developed because RF regression did not have sufficient training for the higher pollutant concentration, therefore the model could not predict the value above the boundary. Besides, the sensor noise of the aircube sensor is ±15 ppb and it has been figured out that NO and NO_2 concentration in the calibration site were in the range that had a high possibility to affected by this noise. The concentration time-series exhibited that the model prediction is really bad when the reference concentration is below 10 ppb.

The sensor performance in the deployment sites was only explored with NO_2. The comparison with the passive sampler data again revealed the same problems that had been already shown during the calibration evaluation. In AC009 and AC010, the over-estimation of the concentration indicated the effect of sensor noise on the prediction, while the estimation from RF models for high NO_2 concentration in AC012 showed the upper limitation on their prediction. Nonetheless, the comparison with the data from continuous monitoring stations in Zurich claimed that the low-cost sensor performed a good prediction of the pollutant concentration in the urban area. The similarity between concentration distributions supported this statement. Lastly, relocation of the aircube sensor to the calibration site explained that the calibration model developed a year ago can be utilized for the model prediction, but the degradation of the performance was clearly shown in the statistic metrics as well as concentration time-series.

In conclusion, the performance of the low-cost sensor in this study shows a good potential

of the technology to be developed as a sophisticated measure to monitor urban air pollution. Though advanced research would be imperative to improve the performance. A recommendation for further study is the comparison between low-cost sensor performance with the pollution modeling data. Since the modeling approach is also highlighted as a good measure for urban air quality monitoring, the comparison can give a good insight into the research field. For instance, the results of the GRAMM/GRAL model in Zurich [11] can be compared to explore whether the low-cost sensors accurately capture spatial and temporal inhomogeneity of the pollutants. GRAMM/GRAL model was already simulated for NOx concentration map of 2013-2014 and the study has proved that the temporal variability of the pollutant was well-documented in various scales [11].

Bibliography

[1] World Health Organization. Regional Office for Europe & European Centre for Environment and Health. (2005). Effects of air pollution on children's health and development : a review of the evidence. Copenhagen : WHO Regional Office for Europe

[2] Latzin, P., Röösli, M., Huss, A., Kuehni, C., & Frey, U. (2009). Air pollution during pregnancy and lung function in newborns: A birth cohort study. The European Respiratory Journal,33(3), 594-603.

[3] Künzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., ... & Schneider, J. (2000). Public-health impact of outdoor and traffic-related air pollution: a European assessment. The Lancet, 356(9232), 795-801.

[4] Pascal, M., Corso, Chanel, Declercq, Badaloni, Cesaroni, . . . On Behalf of the Aphekom Group. (2013). Assessing the public health impacts of urban air pollution in 25 European cities: Results of the Aphekom project. The Science of the Total Environment, 449, 390-400.

[5] Mage, D., Ozolins, G., Peterson, P., Webster, A., Orthofer, R., Vandeweerd, V., & Gwynne, M. (1996). Urban air pollution in megacities of the world. Atmospheric environment, 30(5), 681-686.

[6] Bigi, A., Mueller, M., Grange, S. K., Ghermandi, G., & Hueglin, C. (2018). Performance of NO, NO 2 low cost sensors and three calibration approaches within a real world application.

[7] Zimmerman, N., Presto, A. A., Kumar, S. P., Gu, J., Hauryliuk, A., Robinson, E. S., ... & Subramanian, R. (2018). A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. Atmospheric Measurement Techniques, 11(1), 291-313.

[8] Malings, C., Tanzer, R., Hauryliuk, A., Kumar, S. P., Zimmerman, N., Kara, L. B., ... & Subramanian, R. (2019). Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring. Atmospheric Measurement Techniques, 12(2), 903-920.

[9] Mueller, M., Meyer, J., & Hueglin, C. (2017). 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, 10(10), 3783.

[10] Mueller, M. D., Wagner, M., Barmpadimos, I., & Hueglin, C. (2015). Two-week NO2 maps for the City of Zurich, Switzerland, derived by statistical modelling utilizing

data from a routine passive diffusion sampler network. Atmospheric Environment, 106, 1-10.

[11] Berchet, A., Zink, K., Oettl, D., Brunner, J., Emmenegger, L., & Brunner, D. (2017). Evaluation of high-resolution GRAMM–GRAL (v15. 12/v14. 8) NO x simulations over the city of Zürich, Switzerland. Geoscientific Model Development, 10(9), 3441.

[12] Berchet, A., Zink, K., Muller, C., Oettl, D., Brunner, J., Emmenegger, L., & Brunner, D. (2017). A cost-effective method for simulating city-wide air flow and pollutant dispersion at building resolving scale. Atmospheric environment, 158, 181-196.

[13] Empa Dübendorf, (2020), Technischer Bericht zum Nationalen Beobachtungsnetz für Luftfremdstoffe (NABEL) 2020

[14] Karagulian, F., Barbiere, M., Kotsev, A., Spinelle, L., Gerboles, M., Lagler, F., ... & Borowiak, A. (2019). Review of the performance of low-cost sensors for air quality monitoring. Atmosphere, 10(9), 506.

[15] Schneider, P., Bartonova, A., Castell, N., Dauge, F. R., Gerboles, M., Hagler, G. S., ... & Mijling, B. (2019). Toward a unified terminology of processing levels for low-cost air-quality sensors, Environmental Science & Technology, 53, (15), 8485-8487, DOI: 10.1021/acs.est.9b03950

[16] Christoph Hüglin, (2020), Effect of relocation on the performance of electrochemical sensors for NO_2, Microsenor for air quality workshop, Empa, Laboratory for Air Pollution and Environmental Technology

[17] Jiao, W., Hagler, G., Williams, R., Sharpe, R., Brown, R., Garver, D., ... & Weinstock, L. (2016). Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. Atmospheric Measurement Techniques, 9(11).

[18] Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M., & Bonavitacola, F. (2017). 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, 238, 706-715.

[19] Andersen, R. (2008). Modern methods for robust regression (No. 152). Sage.

[20] Brian Ripley, (2020), Support Functions and Datasets for Venables and Ripley's MASS, Ver. 7.3-53, http://www.stats.ox.ac.uk/pub/MASS4/, Access date : 04.12.2020.

[21] Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.

[22] Andy Liaw, (2018), Breiman and Cutler's Random Forests for Classification and Regression, Ver. 4.6-14, https://www.stat.berkeley.edu/~breiman/RandomForests/, Access date : 04.12.2020.

[23] Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), e1301.

[24] Fushiki, T. (2011). Estimation of prediction error by using K-fold cross-validation. Statistics and Computing, 21(2), 137-146.

[25] Rodriguez, J. D., Perez, A., & Lozano, J. A. (2009). Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE transactions on pattern analysis and machine intelligence, 32(3), 569-575.

[26] Oscar Perpinan Lamigueiro, (2018), Package 'tdr' : Target Diagram, Ver. 0.13, http://github.com/oscarperpinan/tdr, Access date : 08.12.2020.

[27] Felicity Boyd Enders, (2020), Coefficient of determination, Encyclopædia Britannica, https://www.britannica.com/science/coefficient-of-determination, Access date : 08.12.2020.

[28] Mueller, M., Meyer, J., and Hueglin, C., (2017), Design of an ozone and nitrogen dioxide sensor unit and its long-term operation within a sensor network in the city of Zurich, Atmos. Meas. Tech., 10, 3783–3799, https://doi.org/10.5194/amt-10-3783-2017.

[29] *Decentlab GmbH*, Gas sensor datasheets for the aircube sensor unit, the related document available as a link in https://www.decentlab.com/products/air-quality-station-no2-no-co-ox-for-lorawan

Appendices

A Information on site and data

A.1 Site locations on the map

A.1.1 Harkingen Monitoring station

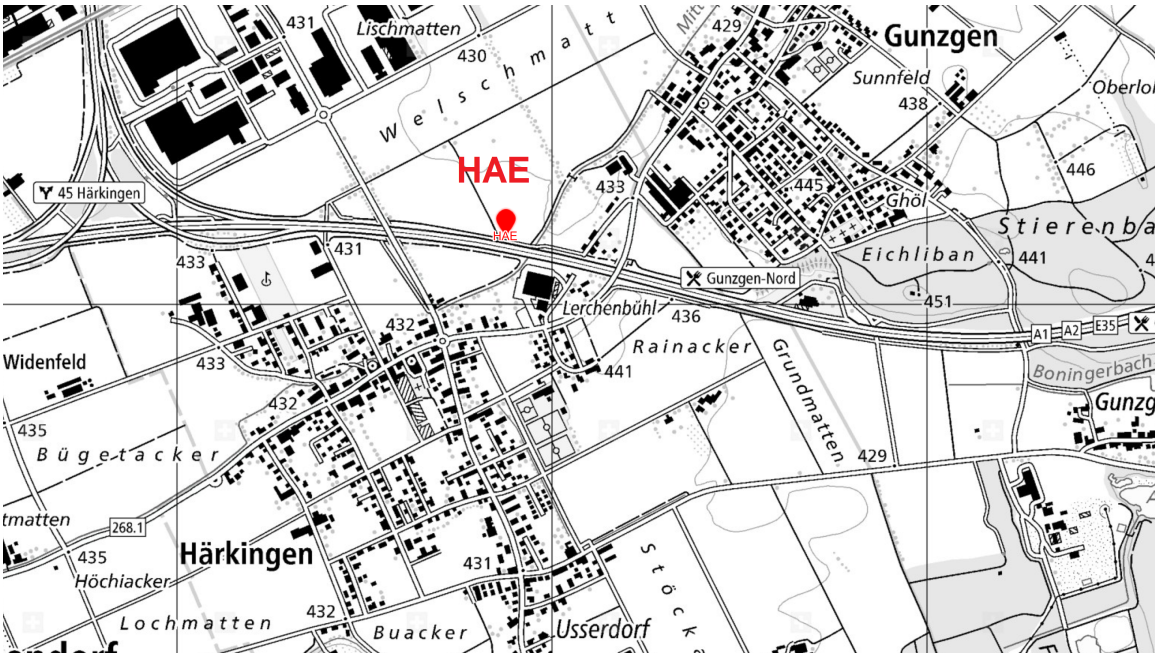


Figure 11: Geographical location of the calibration site, Harkingen air quality monitoring station. The map is obtained from the geo-mapping platform of the Swiss Confederation. `https://map.geo.admin.ch`

(a) View to the south-west side

(b) View to the north-east side

Figure 12: Street view of Harkingen monitoring station. The images are referred to the workshop presentation from Empa. [16]

A.1.2 Deployment sites

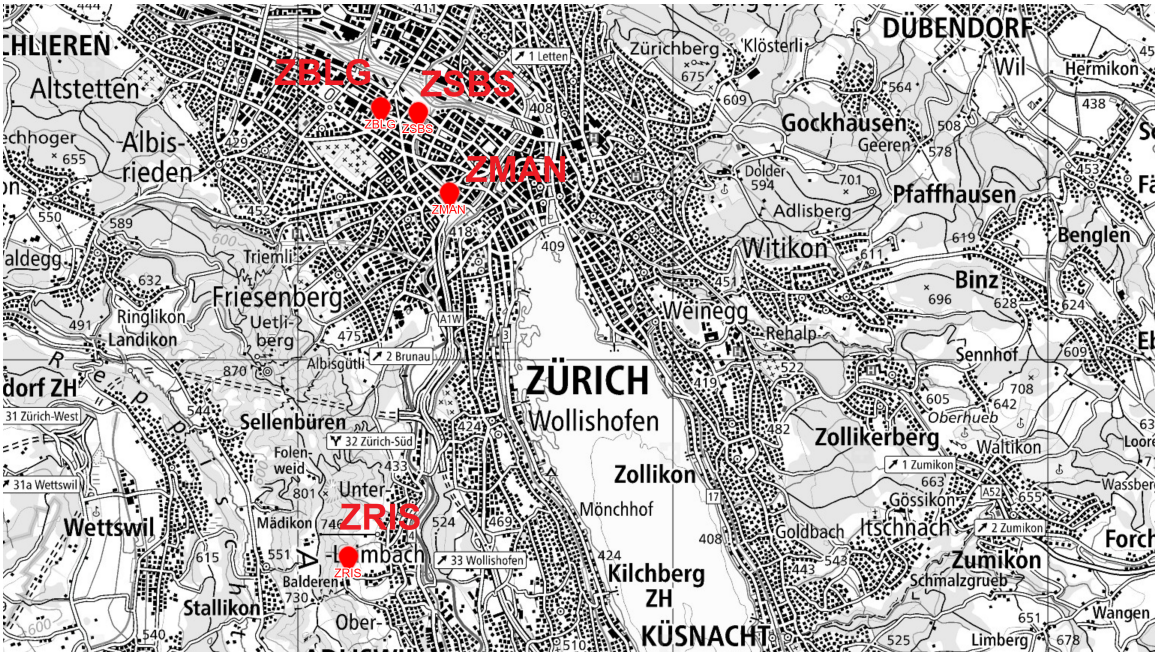


Figure 13: Geographical location of aircube sensor deployment sites in Zurich. The map is obtained from the geo-mapping platform of the Swiss Confederation. `https://map.geo.admin.ch`

(a) ZRIS (AC009)

(b) ZSBS (AC011)

(c) ZBLG (AC010)

(d) ZMAN (AC012)

Figure 14: Street views of four sensor deployment sites in Zurich. The image of ZRIS, ZSBS, and ZMAN were provided from Empa and that of ZBLG was taken from the google street view. `https://goo.gl/maps/DiyXEUgaC6ND3xUb8`

A.1.3 Zurich monitoring stations

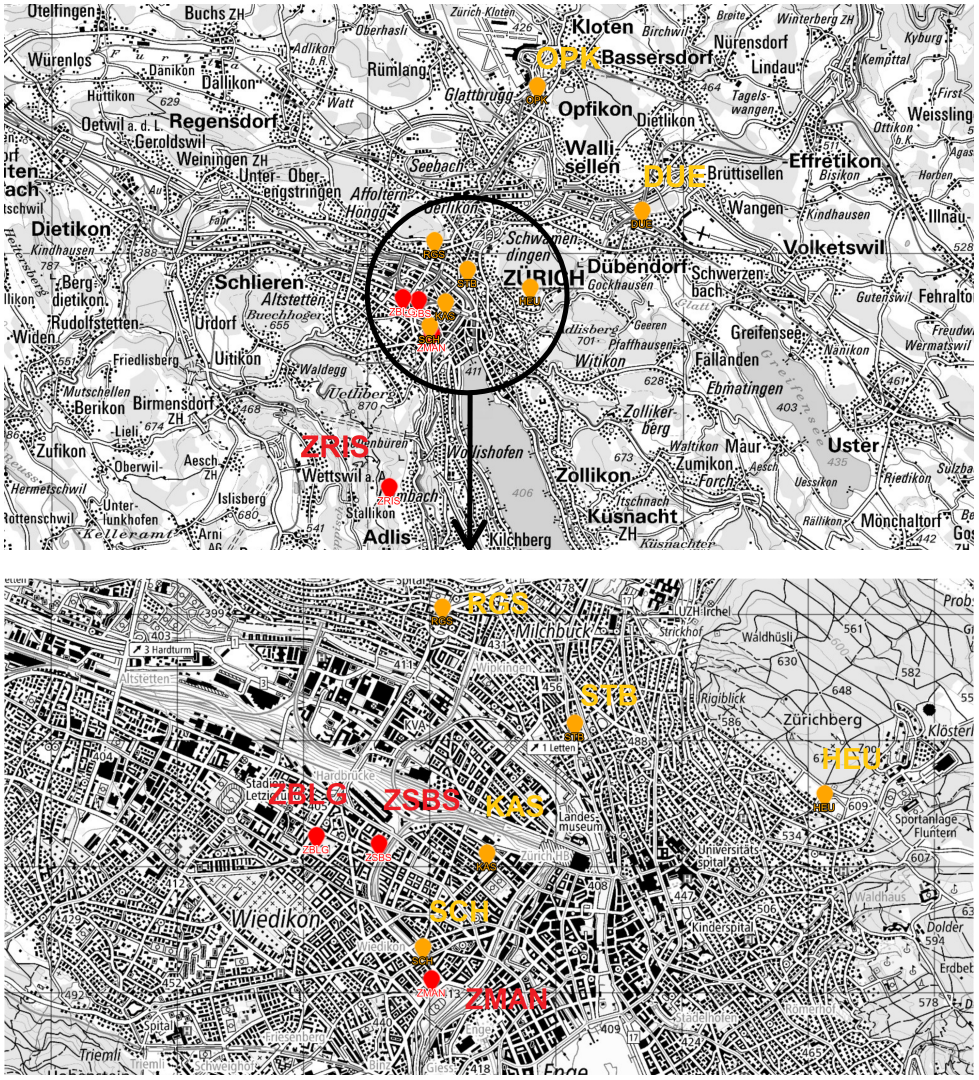


Figure 15: Geographical location of continuous monitoring stations in Zurich with the deployment sites. The city center was zoomed-in at the second map. The map is obtained from the geo-mapping platform of the Swiss Confederation. `https://map.geo.admin.ch`

Table 7: The coordinate system of the continuous monitoring stations in Zurich and their temporal resoluation for NO and NO_2 concentration. DUE and KAS are operated by NABEL, whereas the other stations are operated by OSTLUFT.

Location name (Abbreviation)	Coordinate system (LV95)	Data available from	Data available until	Resolution
Duebendorf (DUE)	1,250,900 N 2,688,675 E	01. Nov. 2018	31. Dec. 2020	10 min
Kaserne (KAS)	1,247,990 N 2,682,450 E	01. Nov. 2018	31. Dec. 2020	10 min
Schimmelstrasse (SCH)	1,247,245 N 2,681,942 E	01. Jan. 2018	30. Dec. 2019	30 min
Heubeeribüel (HEU)	1,248,460 N 2,685,126 E	01. Nov. 2018	30. Dec. 2019	30 min
Stampfenbachstrasse (STB)	1,249,020 N 2,683,146 E	01. Nov. 2018	30. Dec. 2019	30 min
Rosengartenstrasse (RGS)	1,249,940 N 2,682,096 E	01. Nov. 2018	30. Dec. 2019	30 min
Opfikon-Balsberg (OPK)	1,254,830 N 2,685,351 E	01. Nov. 2018	30. Dec. 2019	60 min

A.2 Aircube sensor unit of *Decentlab GmbH*

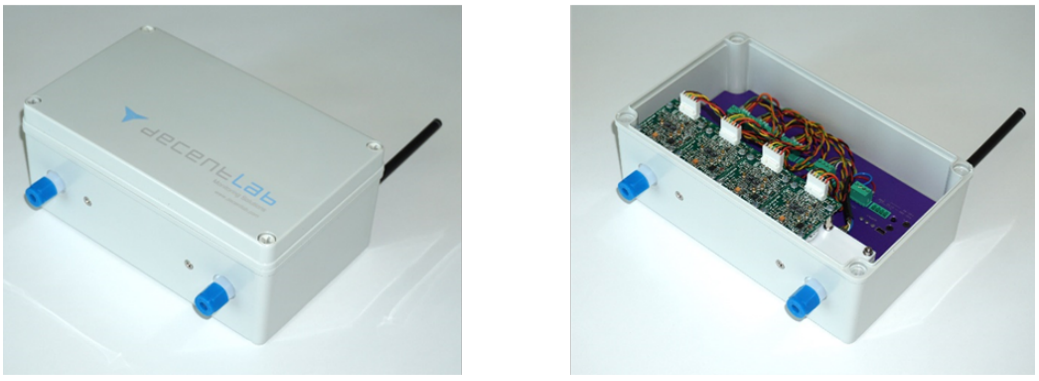


Figure 16: External and internal image of aircube sensor unit. [16]

A device of the aircube sensor utilized in this study can contain maximum four electrochemical gas sensors of NO, NO_2, CO, H_2S, O_3+NO_2, and SO_2. The name of the gas sensors with their range and noise are shown in table 8. Ambient air enters the device through a teflon (PTFE) manifold to which these gas sensors are connected [6]. In addition, the sensor of temperature and relative humidity are included in a unit. The data transmission module uses GSM/GPRS connection [6].

Table 8: Information on the gas sensors for the aircube sensor device. The information is provided in the company website.
`https://www.decentlab.com/products/air-quality-station-no2-no-co-ox-for-lorawan`

Gas	Sensor	Range [ppm]	Noise [ppb]
NO	Alphasense NO-B4	0 - 20	± 15
NO_2	Alphasense NO2-B43F	0 - 20	± 15
CO	Alphasense CO-B4	0 - 1000	± 4
H_2S	Alphasense H2S-B4	0 - 100	± 15
O_3+NO_2	Alphasense OX-B431	0 - 20	± 15
SO_2	Alphasense SO2-B4	0 - 100	± 5
Temperature	-	-40 - +125 °C	± 0.1 - 2 °C
Relative humidity	-	0 - 100 %	± 1.5 - 2 %

B Calibration model formulas

Table 9: Model formula of calibration models using robust linear regression. For the <Reference>, NO and NO_2 concentration from Harkingen monitoring station are assigned. <V_{sensor}> indicates the four sensor data (two NO sensors and two NO_2 sensors). T and RH parameter represents the temperature and relative humidity information respectively.

Model number	Model formula
RLM_1	Reference = function(offset of V_{sensor})
RLM_2	Reference = function(V_{sensor})
RLM_3	Reference = function(V_{sensor}, T)
RLM_4	Reference = function(V_{sensor}, RH)
RLM_5	Reference = function(V_{sensor}, T, RH)
RLM_6	Reference = function(V_{sensor}, T, D_{RH_60})
RLM_7	Reference = function(V_{sensor}, T, D_{RH_90})
RLM_8	Reference = function(V_{sensor}, T, D_{RH_120})
RLM_9	Reference = function(V_{sensor}, T, D_{RH_150})
RLM_10	Reference = function(V_{sensor}, RH, D_{RH_60})
RLM_11	Reference = function(V_{sensor}, RH, D_{RH_90})
RLM_12	Reference = function(V_{sensor}, RH, D_{RH_120})
RLM_13	Reference = function(V_{sensor}, RH, D_{RH_150})
RLM_14	Reference = function(V_{sensor}, T, RH, D_{RH_60})
RLM_15	Reference = function(V_{sensor}, T, RH, D_{RH_90})
RLM_16	Reference = function(V_{sensor}, T, RH, D_{RH_120})
RLM_17	Reference = function(V_{sensor}, T, RH, D_{RH_150})
RLM_18	Reference = function(V_{sensor}, T, V_{sensor}*T)
RLM_19	Reference = function(V_{sensor}, T, V_{sensor}*T, D_{RH_60})
RLM_20	Reference = function(V_{sensor}, T, V_{sensor}*T, D_{RH_90})
RLM_21	Reference = function(V_{sensor}, T, V_{sensor}*T, D_{RH_120})
RLM_22	Reference = function(V_{sensor}, T, V_{sensor}*T, D_{RH_150})
RLM_23	Reference = function(V_{sensor}, T, V_{sensor}*T, V_{sensor}*T^2, D_{RH_60})
RLM_24	Reference = function(V_{sensor}, T, V_{sensor}*T, V_{sensor}*T^2, D_{RH_90})
RLM_25	Reference = function(V_{sensor}, T, V_{sensor}*T, V_{sensor}*T^2, D_{RH_120})
RLM_26	Reference = function(V_{sensor}, T, V_{sensor}*T, V_{sensor}*T^2, D_{RH_150})

Table 10: The model formula of calibration models using random forest regression. For <Reference>, NO and NO_2 concentration from the Harkingen monitoring station are assigned. V_{sensor} indicates the four sensor data (two NO sensors and two NO_2 sensors). T and RH parameter represents the temperature and relative humidity information respectively.

Model number	Model formula
RF_1	Reference = function(V_{sensor})
RF_2	Reference = function(V_{sensor}, T)
RF_3	Reference = function(V_{sensor}, T, D_{RH_60})
RF_4	Reference = function(V_{sensor}, T, D_{RH_90})
RF_5	Reference = function(V_{sensor}, T, D_{RH_120})
RF_6	Reference = function(V_{sensor}, T, D_{RH_150})
RF_7	Reference = function(V_{sensor}, RH)
RF_8	Reference = function(V_{sensor}, T, RH)
RF_9	Reference = function(V_{sensor}, T, V_{sensor}*T)

C Statistical metrics in the target diagram

Table 11: Statistical metrics in the target diagram. The table is adopted from the supplemantary materials of Zimmerman, N. et al (2018) [7]. In the formula, M represents the measured concentration from the low-cost sensor and O represents the observation value from the reference equipment.

Metrics	Abbreviation	Formula
Root mean square error	RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - O_i)^2}$
Normalized root mean square error	nRMSE	$\frac{RMSE}{\sigma_O}$
Mean bias error	MBE	$\overline{M} - \overline{O}$
Centerd root mean square error	CRMSE	$\sqrt{RMSE^2 - MBE^2}$

Table 11 shows four metrics that can be estimated from the target diagram. RMSE is a non-negative measure that depicts the difference between the measured value and predicted value, and it is sensitive to extreme values [7]. nRMSE and CRMSE are corrected versions of RMSE; nRMSE is normalized by the standard deviation of reference values, whereas the CRMSE is a bias-corrected RMSE that neglects MBE from RMSE. MBE is a

measure that represents an averaged difference between measured values and estimation in whole periods.

D Results of sensor calibration

D.1 Node size decision

Table 12: The result of node size experiments in NO_00 sensors of AC009. nRMSE values are arranged by each node size and each random forest regression models.

Node size	RF_1	RF_2	RF_3	RF_4	RF_5	RF_6	RF_7	RF_8	RF_9
5	0.665	0.482	0.456	0.457	0.455	0.454	0.541	0.460	0.482
10	0.652	0.481	0.457	0.458	0.454	0.454	0.535	0.458	0.478
20	0.632	0.476	0.454	0.455	0.452	0.453	0.530	0.458	0.475
30	0.623	0.472	0.454	0.455	0.456	0.452	0.525	0.458	0.472
40	0.618	0.472	0.453	0.456	0.453	0.453	0.522	0.458	0.472
50	0.614	0.470	0.455	0.457	0.454	0.453	0.522	0.460	0.470
100	0.605	0.467	0.459	0.459	0.458	0.457	0.518	0.463	0.468
200	0.598	0.469	0.466	0.468	0.466	0.467	0.519	0.470	0.467
300	0.595	0.469	0.473	0.474	0.470	0.472	0.52	0.477	0.469

Table 13: The result of node size experiments in NO_01 sensors of AC010. nRMSE values are arranged by each node size and each random forest regression models.

Node size	RF_1	RF_2	RF_3	RF_4	RF_5	RF_6	RF_7	RF_8	RF_9
5	0.793	0.500	0.469	0.463	0.463	0.464	0.620	0.469	0.500
10	0.774	0.494	0.466	0.461	0.463	0.461	0.615	0.469	0.493
20	0.754	0.488	0.465	0.461	0.462	0.462	0.607	0.467	0.490
30	0.743	0.484	0.465	0.460	0.461	0.461	0.603	0.467	0.485
40	0.736	0.483	0.462	0.462	0.462	0.461	0.601	0.467	0.482
50	0.733	0.481	0.465	0.464	0.464	0.462	0.601	0.466	0.481
100	0.722	0.479	0.469	0.467	0.469	0.469	0.597	0.471	0.479
200	0.715	0.480	0.478	0.481	0.478	0.477	0.597	0.483	0.480
300	0.713	0.482	0.487	0.488	0.49	0.486	0.598	0.491	0.483

Table 14: The result of node size experiments in NO2_00 sensors of AC011. nRMSE values are arranged by each node size and each random forest regression models.

Node size	RF_1	RF_2	RF_3	RF_4	RF_5	RF_6	RF_7	RF_8	RF_9
5	0.889	0.822	0.690	0.666	0.652	0.640	0.839	0.744	0.824
10	0.877	0.813	0.688	0.666	0.650	0.639	0.829	0.741	0.814
20	0.862	0.802	0.686	0.667	0.649	0.640	0.816	0.742	0.803
30	0.855	0.797	0.687	0.665	0.651	0.638	0.812	0.744	0.796
40	0.849	0.794	0.688	0.666	0.652	0.642	0.808	0.747	0.796
50	0.846	0.793	0.690	0.669	0.656	0.641	0.807	0.747	0.791
100	0.837	0.789	0.695	0.676	0.662	0.649	0.801	0.754	0.789
200	0.830	0.790	0.708	0.690	0.674	0.665	0.801	0.767	0.790
300	0.827	0.793	0.715	0.702	0.688	0.678	0.802	0.774	0.792

Table 15: The result of node size experiments in NO2_01 sensors of AC012. nRMSE values are arranged by each node size and each random forest regression models.

Node size	RF_1	RF_2	RF_3	RF_4	RF_5	RF_6	RF_7	RF_8	RF_9
5	0.719	0.645	0.540	0.533	0.529	0.526	0.672	0.599	0.642
10	0.712	0.637	0.536	0.532	0.528	0.523	0.661	0.597	0.637
20	0.701	0.630	0.537	0.529	0.528	0.523	0.652	0.596	0.627
30	0.695	0.626	0.536	0.529	0.527	0.524	0.646	0.597	0.625
40	0.690	0.623	0.537	0.532	0.527	0.523	0.643	0.599	0.622
50	0.687	0.621	0.537	0.532	0.527	0.524	0.640	0.600	0.622
100	0.678	0.619	0.542	0.535	0.533	0.529	0.636	0.604	0.619
200	0.672	0.621	0.553	0.549	0.542	0.543	0.637	0.615	0.621
300	0.669	0.623	0.564	0.559	0.555	0.555	0.637	0.619	0.623

D.2 Model selection

Table 16: The result of model selection in AC009 and AC010. nRMSE values in each models of the sensors are presented in the table.

Sensor / Model	AC009				AC010			
	NO_00	NO_01	NO2_00	NO2_01	NO_00	NO_01	NO2_00	NO2_01
RLM_1	0.719	0.731	0.841	0.830	0.754	0.784	0.829	0.806
RLM_2	0.646	0.684	0.827	0.761	0.691	0.751	0.822	0.733
RLM_3	0.607	0.640	0.769	0.699	0.612	0.632	0.766	0.683
RLM_4	0.629	0.670	0.766	0.718	0.661	0.700	0.773	0.699
RLM_5	0.597	0.619	0.763	0.699	0.594	0.611	0.762	0.683
RLM_6	0.606	0.638	0.731	0.635	0.610	0.631	0.716	0.625
RLM_7	0.605	0.635	0.719	0.630	0.608	0.628	0.702	0.619
RLM_8	0.603	0.631	0.709	0.629	0.606	0.625	0.691	0.615
RLM_9	0.601	0.627	0.700	0.630	0.603	0.622	0.682	0.613
RLM_10	0.628	0.668	0.727	0.662	0.660	0.698	0.724	0.645
RLM_11	0.627	0.666	0.719	0.659	0.658	0.696	0.714	0.641
RLM_12	0.626	0.663	0.712	0.658	0.656	0.693	0.706	0.638
RLM_13	0.624	0.660	0.707	0.660	0.653	0.690	0.701	0.637
RLM_14	0.597	0.619	0.724	0.635	0.593	0.611	0.711	0.625
RLM_15	0.597	0.618	0.714	0.628	0.593	0.610	0.699	0.618
RLM_16	0.596	0.616	0.705	0.625	0.592	0.609	0.689	0.613
RLM_17	0.595	0.614	0.698	0.624	0.591	0.608	0.681	0.610
RLM_18	0.496	0.508	0.713	0.666	0.499	0.515	0.718	0.661
RLM_19	0.492	0.500	0.661	0.599	0.493	0.508	0.659	0.603
RLM_20	0.491	0.498	0.648	0.596	0.491	0.505	0.645	0.598
RLM_21	0.490	0.496	0.637	0.597	0.489	0.502	0.635	0.595
RLM_22	0.489	0.494	0.628	0.600	0.487	0.499	0.628	0.595
RLM_23	0.492	0.500	0.661	0.599	0.493	0.508	0.659	0.603
RLM_24	0.491	0.498	0.648	0.596	0.491	0.505	0.645	0.598
RLM_25	0.490	0.496	0.637	0.597	0.489	0.502	0.635	0.595
RLM_26	0.489	0.494	0.628	0.600	0.487	0.499	0.628	0.595
RF_1	0.605	0.635	0.735	0.702	0.656	0.722	0.744	0.689
RF_2	0.467	0.477	0.665	0.647	0.470	0.478	0.680	0.631
RF_3	0.458	0.466	0.582	0.555	0.460	0.468	0.591	0.551
RF_4	0.458	0.466	0.573	0.553	0.460	0.467	0.583	0.546
RF_5	0.457	0.464	0.567	0.554	0.461	0.468	0.576	0.545
RF_6	0.457	0.464	0.562	0.556	0.460	0.468	0.572	0.546
RF_7	0.519	0.543	0.679	0.664	0.563	0.597	0.700	0.650
RF_8	0.462	0.470	0.646	0.634	0.465	0.470	0.662	0.617
RF_9	0.468	0.477	0.666	0.647	0.470	0.479	0.680	0.631

Table 17: The result of model selection in AC011 and AC012. nRMSE values in each models of the sensors are presented in the table.

Sensor / Model	AC011				AC012			
	NO_00	NO_01	NO2_00	NO2_01	NO_00	NO_01	NO2_00	NO2_01
RLM_1	0.764	0.747	0.909	0.822	0.812	0.768	0.856	0.811
RLM_2	0.705	0.686	0.908	0.788	0.773	0.728	0.851	0.728
RLM_3	0.589	0.636	0.883	0.746	0.572	0.657	0.808	0.677
RLM_4	0.645	0.670	0.893	0.754	0.646	0.704	0.821	0.680
RLM_5	0.578	0.617	0.884	0.745	0.569	0.630	0.808	0.674
RLM_6	0.589	0.635	0.848	0.669	0.572	0.648	0.763	0.630
RLM_7	0.589	0.633	0.835	0.650	0.572	0.644	0.753	0.617
RLM_8	0.588	0.630	0.822	0.636	0.572	0.638	0.744	0.607
RLM_9	0.587	0.627	0.812	0.627	0.572	0.633	0.738	0.599
RLM_10	0.645	0.669	0.857	0.678	0.646	0.698	0.781	0.635
RLM_11	0.644	0.668	0.846	0.664	0.645	0.693	0.774	0.625
RLM_12	0.643	0.666	0.836	0.653	0.644	0.689	0.768	0.617
RLM_13	0.642	0.663	0.828	0.646	0.643	0.684	0.764	0.611
RLM_14	0.578	0.617	0.847	0.667	0.569	0.626	0.763	0.627
RLM_15	0.578	0.617	0.835	0.650	0.569	0.624	0.754	0.615
RLM_16	0.578	0.616	0.822	0.636	0.569	0.621	0.745	0.606
RLM_17	0.578	0.615	0.812	0.626	0.569	0.618	0.738	0.599
RLM_18	0.483	0.491	0.826	0.706	0.512	0.542	0.726	0.646
RLM_19	0.480	0.485	0.753	0.619	0.510	0.528	0.640	0.599
RLM_20	0.478	0.483	0.730	0.600	0.509	0.525	0.624	0.588
RLM_21	0.477	0.481	0.711	0.587	0.508	0.522	0.613	0.578
RLM_22	0.475	0.479	0.695	0.580	0.507	0.519	0.606	0.572
RLM_23	0.480	0.485	0.753	0.619	0.510	0.528	0.640	0.599
RLM_24	0.478	0.483	0.730	0.600	0.509	0.525	0.624	0.588
RLM_25	0.477	0.481	0.711	0.587	0.508	0.522	0.613	0.578
RLM_26	0.475	0.479	0.695	0.580	0.507	0.519	0.606	0.572
RF_1	0.663	0.629	0.836	0.733	0.759	0.686	0.744	0.678
RF_2	0.461	0.469	0.788	0.687	0.478	0.499	0.696	0.619
RF_3	0.454	0.462	0.695	0.574	0.470	0.479	0.569	0.542
RF_4	0.453	0.460	0.676	0.560	0.469	0.476	0.561	0.535
RF_5	0.451	0.458	0.661	0.553	0.469	0.476	0.556	0.531
RF_6	0.450	0.455	0.648	0.547	0.467	0.474	0.554	0.529
RF_7	0.529	0.539	0.802	0.703	0.554	0.588	0.707	0.636
RF_8	0.454	0.463	0.755	0.664	0.473	0.488	0.669	0.603
RF_9	0.462	0.470	0.788	0.687	0.478	0.499	0.695	0.619

D.3 Calibration Evaluation

D.3.1 Target diagram

Table 18: nRMSE values from the target diagram of the sensor calibration. The values are presented in each models.

Unit	Sensor	RLM_A	RLM_B	RLM_C	RF_A	RF_B
AC009	NO_00	0.48	0.49	0.48	0.45	0.46
	NO_01	0.49	0.50	0.49	0.46	0.46
	NO2_00	0.58	0.70	0.61	0.88	0.63
	NO2_01	0.58	0.70	0.61	0.88	0.63
AC010	NO_00	0.48	0.49	0.48	0.45	0.46
	NO_01	0.49	0.51	0.49	0.45	0.46
	NO2_00	0.59	0.71	0.61	0.56	0.66
	NO2_01	0.58	0.66	0.59	0.54	0.62
AC011	NO_00	0.49	0.50	0.49	0.45	0.46
	NO_01	0.49	0.50	0.49	0.46	0.46
	NO2_00	0.69	0.82	0.70	0.89	0.76
	NO2_01	0.69	0.82	0.70	0.89	0.76
AC012	NO_00	0.48	0.49	0.48	0.45	0.46
	NO_01	0.49	0.52	0.49	0.46	0.47
	NO2_00	0.58	0.74	0.61	0.86	0.68
	NO2_01	0.58	0.74	0.61	0.86	0.68

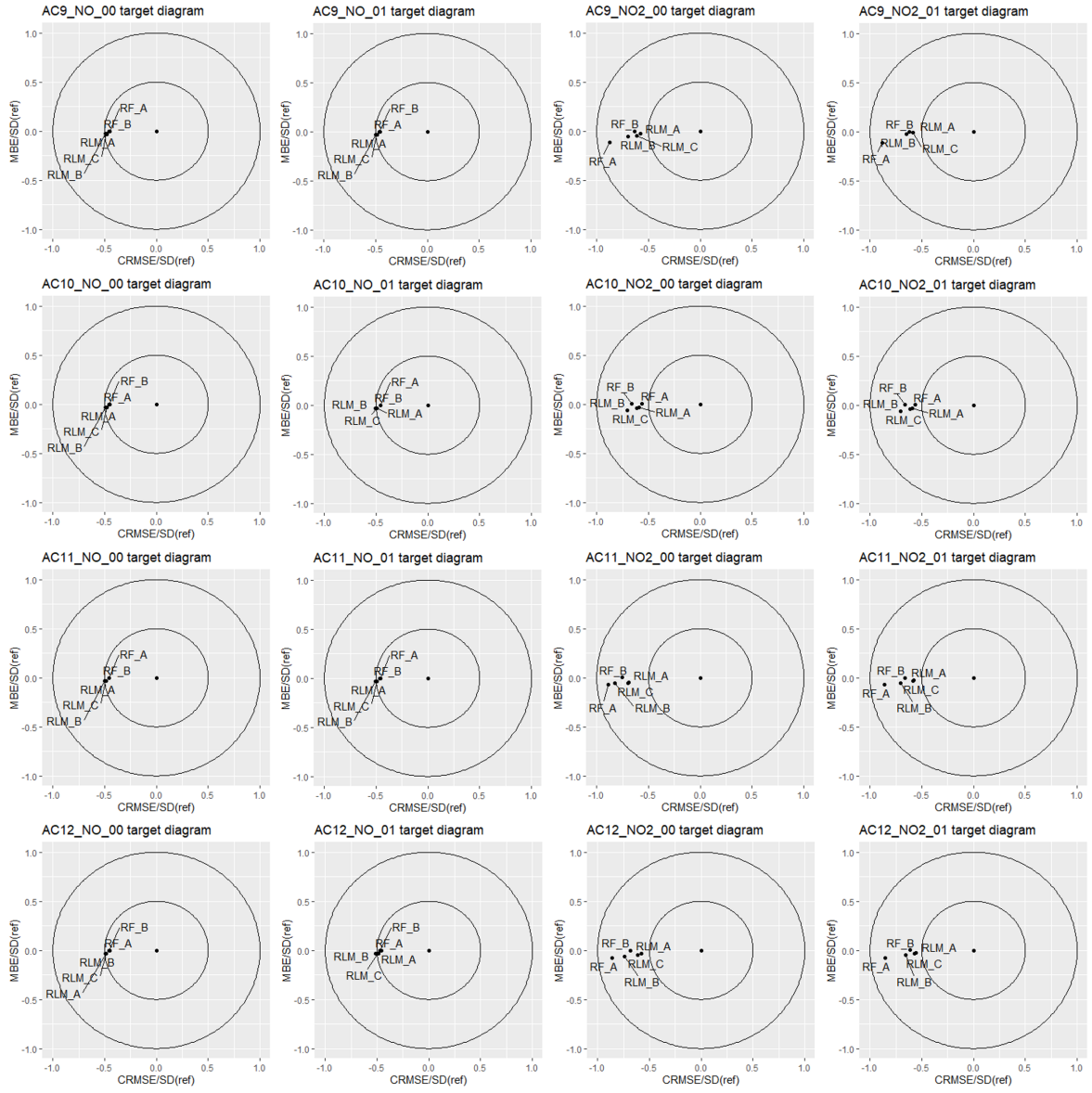


Figure 17: Target diagram of the calibration evaluation from all aircube sensors. AC9, AC10, AC11, and AC12 in the diagram indicates AC009, AC010, AC011, and AC012.

D.3.2 Coefficient of determination

Table 19: Coefficient of determination (R^2) of the sensor calibration. The values are presented in each models.

Unit	Sensor	RLM_A	RLM_B	RLM_C	RF_A	RF_B
AC009	NO_00	0.77	0.76	0.77	0.80	0.79
	NO_01	0.76	0.75	0.76	0.79	0.79
	NO2_00	0.67	0.52	0.63	0.70	0.60
	NO2_01	0.66	0.58	0.66	0.70	0.62
AC010	NO_00	0.77	0.76	0.77	0.80	0.79
	NO_01	0.76	0.74	0.76	0.80	0.79
	NO2_00	0.65	0.50	0.63	0.69	0.57
	NO2_01	0.67	0.57	0.66	0.71	0.62
AC011	NO_00	0.77	0.75	0.76	0.80	0.79
	NO_01	0.76	0.75	0.76	0.79	0.79
	NO2_00	0.52	0.32	0.51	0.58	0.43
	NO2_01	0.67	0.50	0.66	0.71	0.56
AC012	NO_00	0.77	0.76	0.77	0.80	0.80
	NO_01	0.76	0.73	0.76	0.79	0.78
	NO2_00	0.67	0.46	0.63	0.69	0.54
	NO2_01	0.70	0.57	0.67	0.73	0.63

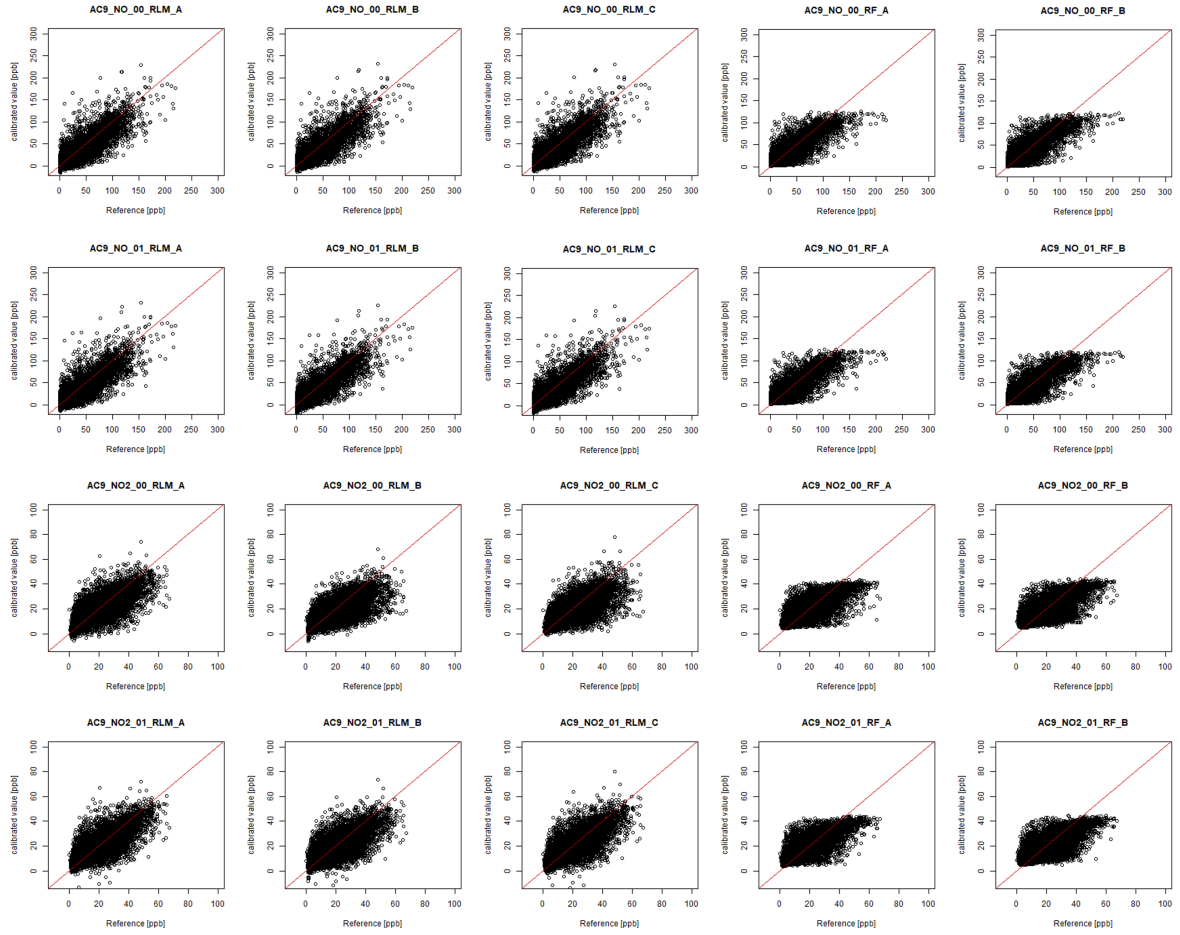


Figure 18: Correlation plots of AC009 for the measured concentration in the calibration site during 29.06.2018 - 12.12.2018 with reference concentration data from the Harkingen monitoring station.

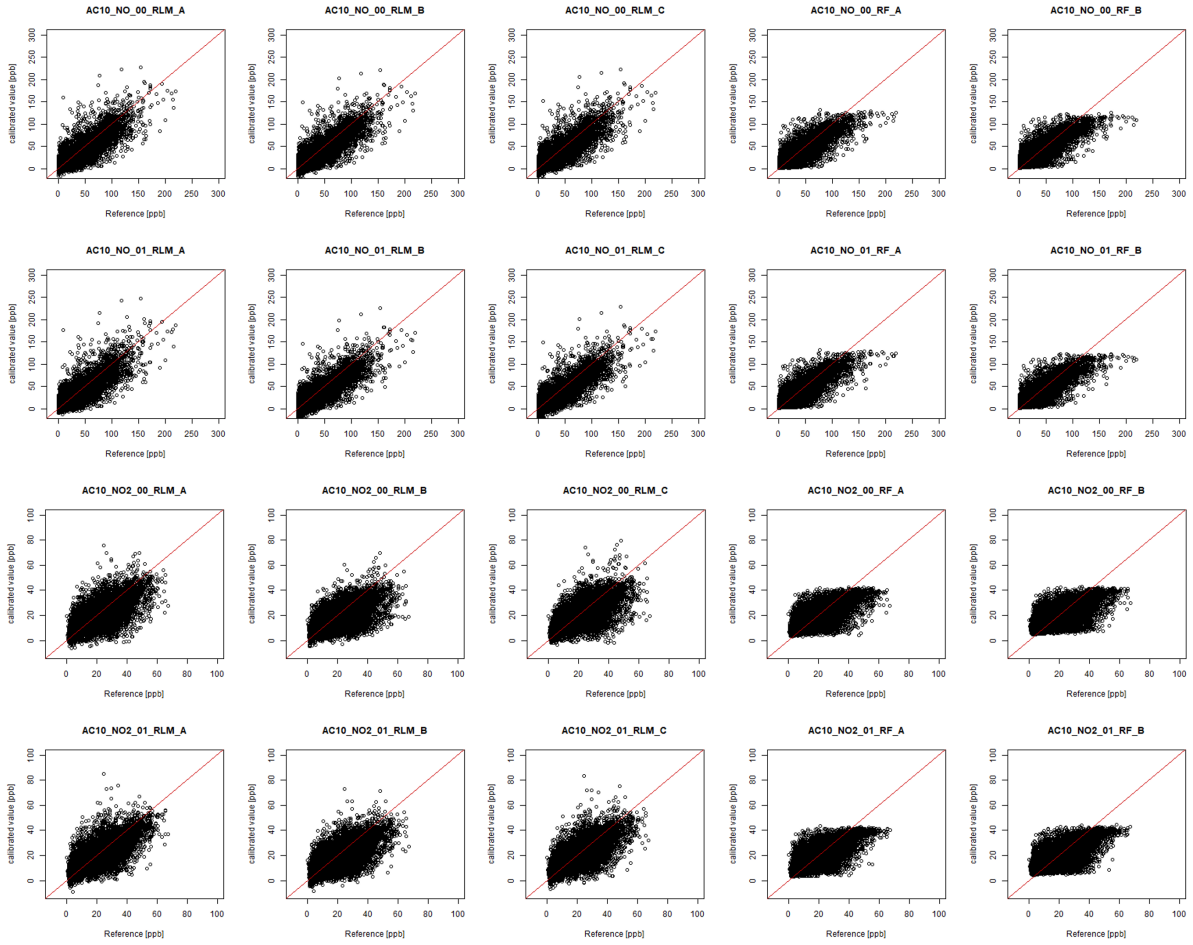


Figure 19: Correlation plots of AC010 for the measured concentration in the calibration site during 29.06.2018 - 12.12.2018 with reference concentration data from the Harkingen monitoring station.

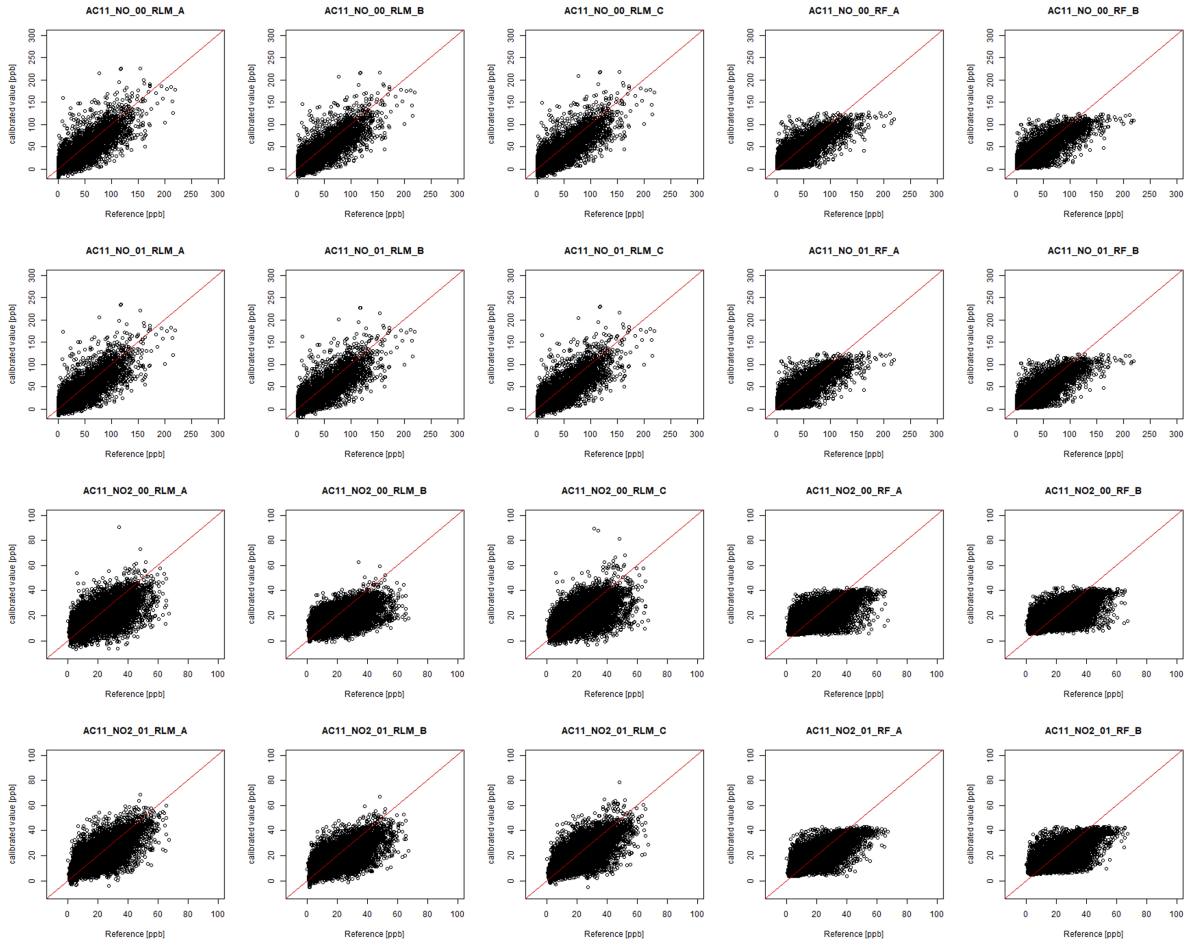


Figure 20: Correlation plots of AC011 for the measured concentration in the calibration site during 29.06.2018 - 12.12.2018 with reference concentration data from the Harkingen monitoring station.

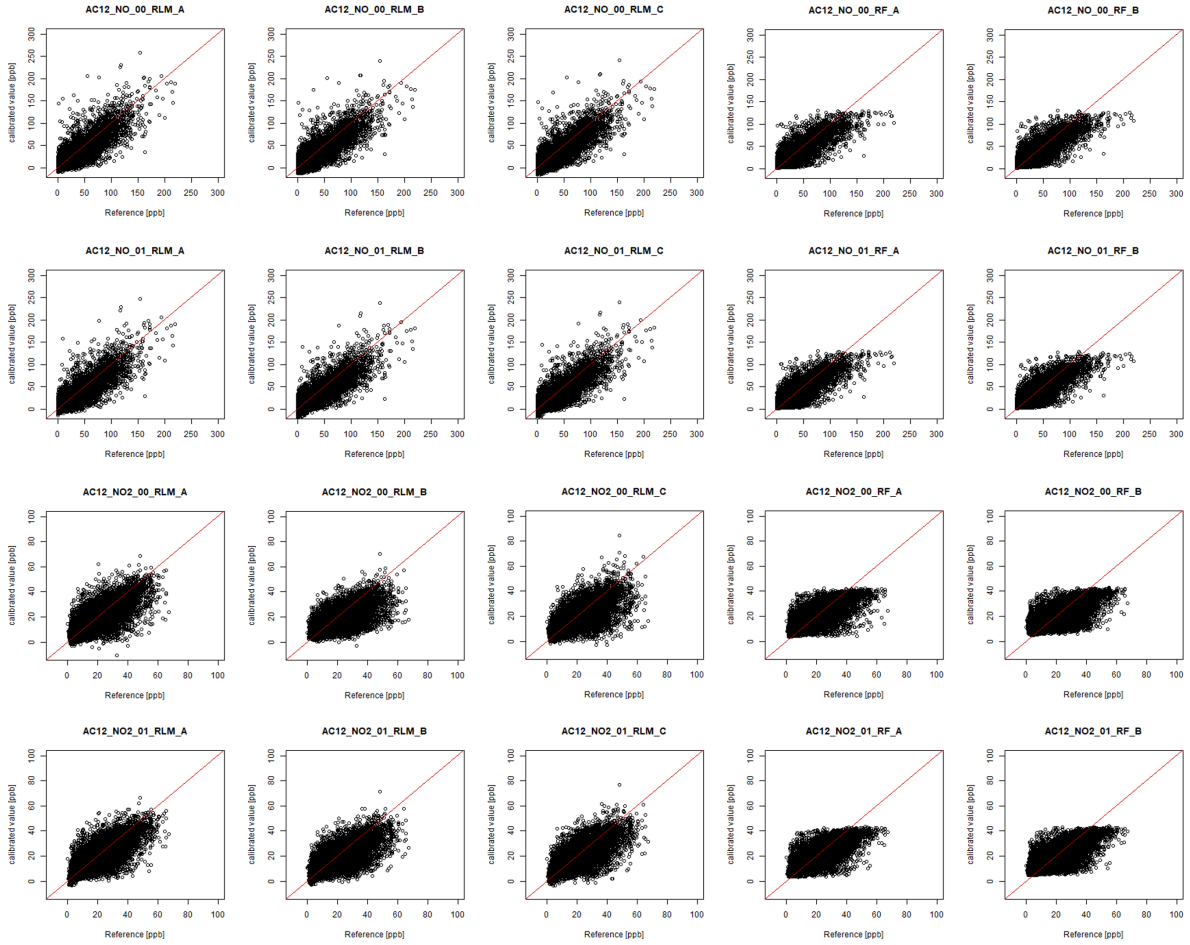


Figure 21: Correlation plots of AC012 for the measured concentration in the calibration site during 29.06.2018 - 12.12.2018 with reference concentration data from the Harkingen monitoring station.

E Results of performance analysis

E.1 Performance analysis in the deployment site

Table 20: Statistic metrics in the comparison of sensor data with the passive sampler data in deployment sites.

Sensor		nRMSE	R-sqaure
AC009	NO2_00	1.06	0.77
	NO2_01	1.46	0.22
AC010	NO2_00	0.32	0.80
	NO2_01	0.34	0.82
AC011	NO2_00	0.24	0.37
	NO2_01	0.12	0.80
AC012	NO2_00	0.64	0.20
	NO2_01	0.44	0.57

E.2 Performance analysis in the calibration site

E.2.1 Target diagram

Table 21: nRMSE values from the target diagram of the sensor performance in the calibration site. The values are presented in each models.

Unit	Sensor	RLM_A	RLM_B	RLM_C	RF_A	RF_B
AC009	NO_00	0.44	0.49	0.45	0.44	0.44
	NO_01	0.47	0.54	0.45	0.45	0.45
	NO2_00	0.70	0.77	0.69	0.87	0.66
	NO2_01	0.76	0.74	0.77	0.92	0.72
AC010	NO_00	0.54	0.54	0.57	0.49	0.50
	NO_01	0.53	0.58	0.48	0.48	0.49
	NO2_00	0.83	0.87	0.85	0.90	0.78
	NO2_01	0.78	0.82	0.81	0.90	0.71
AC011	NO_00	0.50	0.51	0.46	0.46	0.44
	NO_01	0.50	0.55	0.48	0.47	0.45
	NO2_00	0.91	0.93	0.91	0.92	0.78
	NO2_01	0.81	0.81	0.82	0.90	0.76
AC012	NO_00	0.50	0.52	0.45	0.46	0.44
	NO_01	0.50	0.58	0.47	0.46	0.45
	NO2_00	0.75	0.82	0.77	0.87	0.69
	NO2_01	0.67	0.72	0.69	0.86	0.64

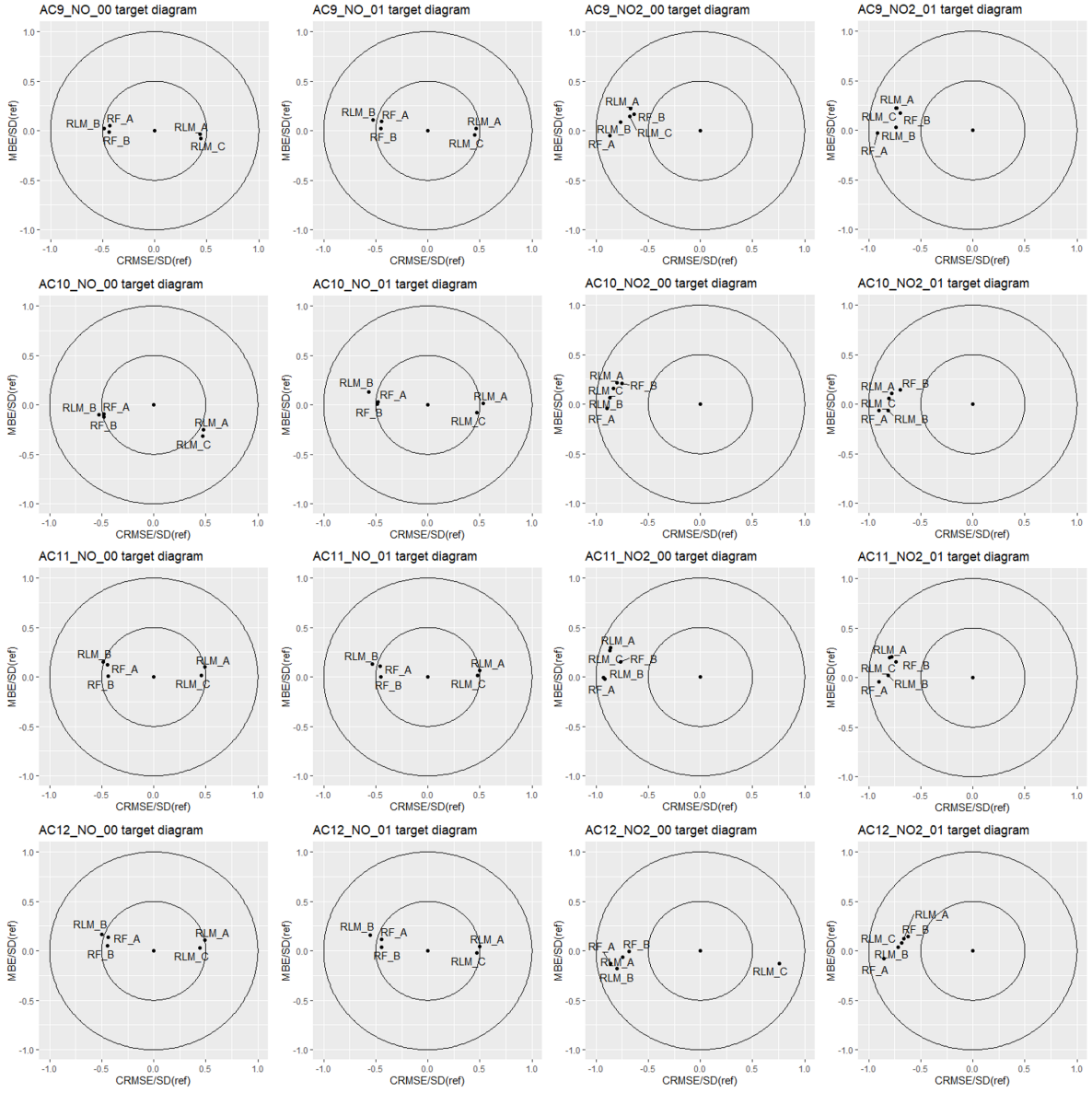


Figure 22: Target diagram of the sensor performance in the calibration site. AC9, AC10, AC11, and AC12 in the diagram indicates AC009, AC010, AC011, and AC012.

E.2.2 Coefficient of determination

Table 22: Coefficient of determination (R^2) of the sensor performance in the calibration site. The values are presented in each models.

Unit	Sensor	RLM_A	RLM_B	RLM_C	RF_A	RF_B
AC009	NO_00	0.83	0.79	0.83	0.81	0.81
	NO_01	0.82	0.75	0.82	0.80	0.80
	NO2_00	0.58	0.42	0.57	0.33	0.60
	NO2_01	0.48	0.45	0.49	0.18	0.52
AC010	NO_00	0.81	0.74	0.80	0.77	0.77
	NO_01	0.79	0.69	0.80	0.77	0.76
	NO2_00	0.40	0.27	0.39	0.24	0.44
	NO2_01	0.41	0.35	0.39	0.23	0.51
AC011	NO_00	0.81	0.77	0.82	0.81	0.81
	NO_01	0.80	0.74	0.81	0.79	0.80
	NO2_00	0.34	0.16	0.35	0.20	0.41
	NO2_01	0.42	0.35	0.43	0.23	0.46
AC012	NO_00	0.81	0.76	0.82	0.81	0.80
	NO_01	0.81	0.71	0.81	0.80	0.80
	NO2_00	0.49	0.36	0.52	0.34	0.53
	NO2_01	0.58	0.49	0.57	0.35	0.62

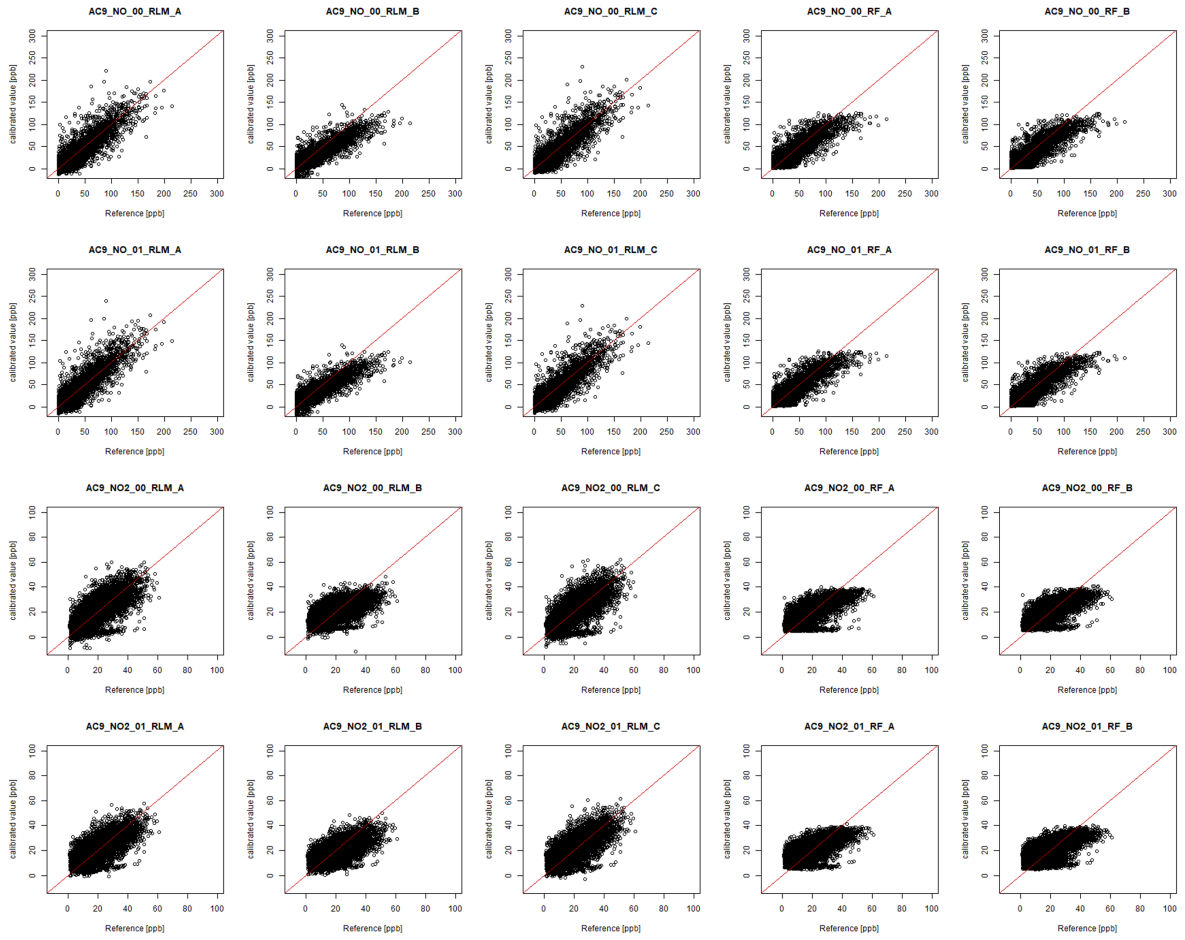


Figure 23: Correlation plots of AC009 for the measured concentration in the calibration site during 12.12.2019 - 31.03.20 with reference concentration data from the Harkingen monitoring station.

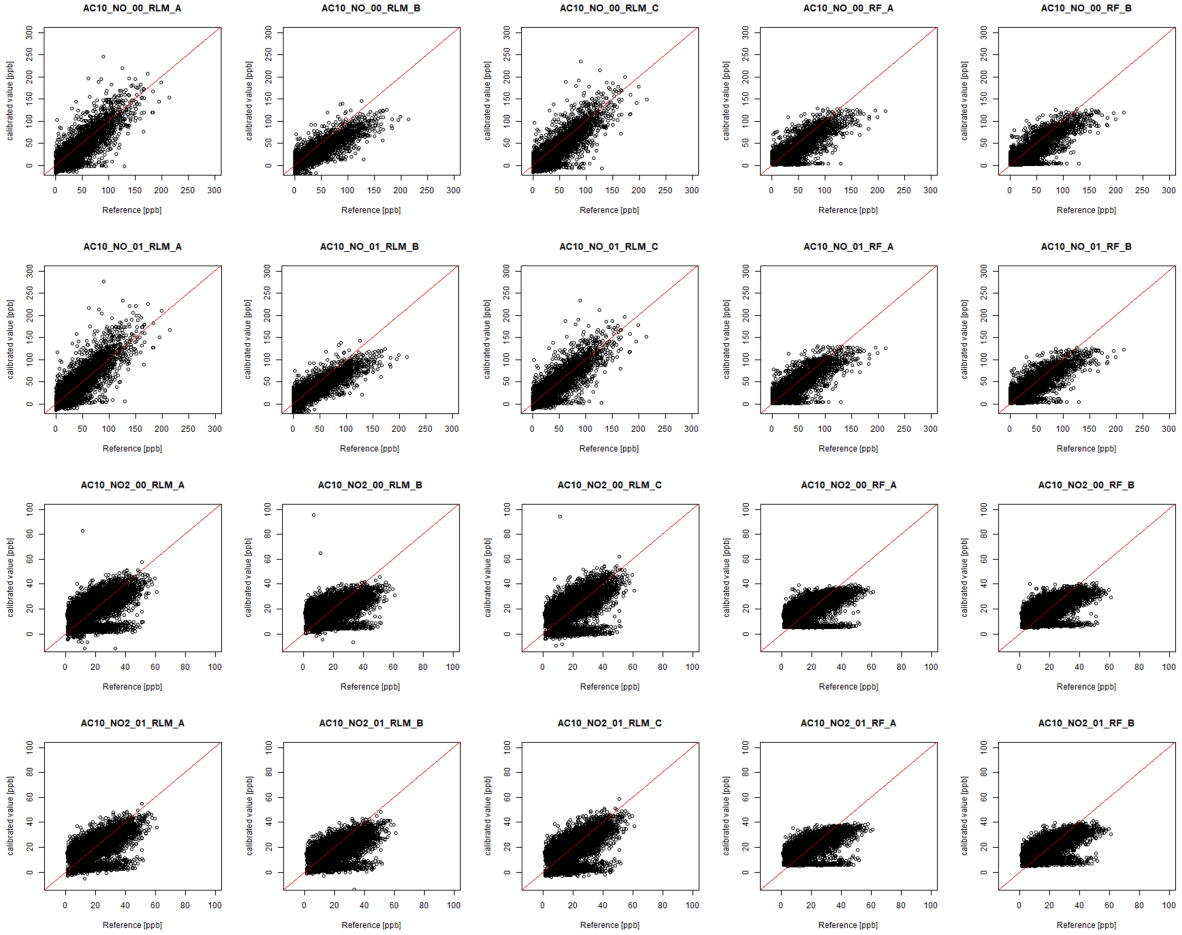


Figure 24: Correlation plots of AC010 for the measured concentration in the calibration site during 12.12.2019 - 31.03.20 with reference concentration data from the Harkingen monitoring station.

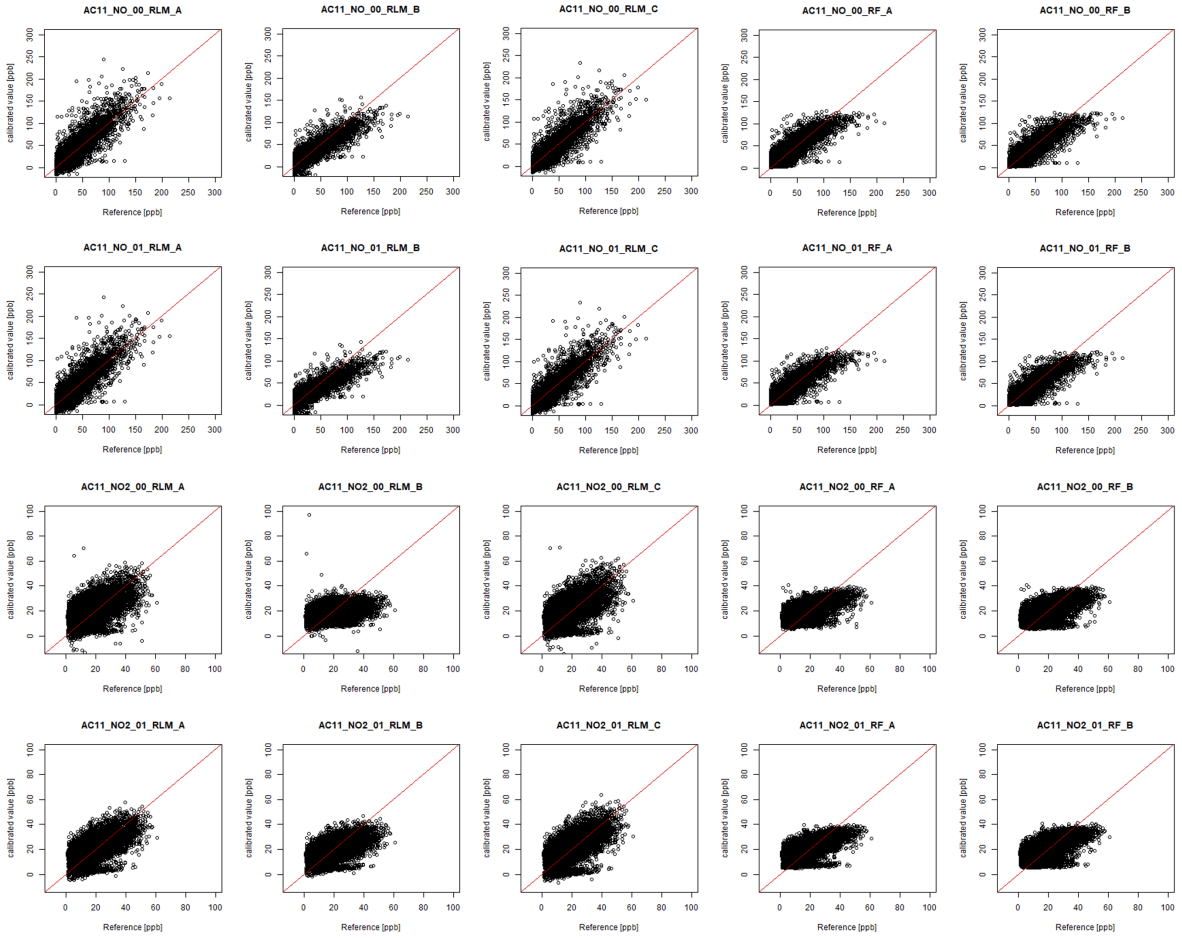


Figure 25: Correlation plots of AC011 for the measured concentration in the calibration site during 12.12.2019 - 31.03.20 with reference concentration data from the Harkingen monitoring station.

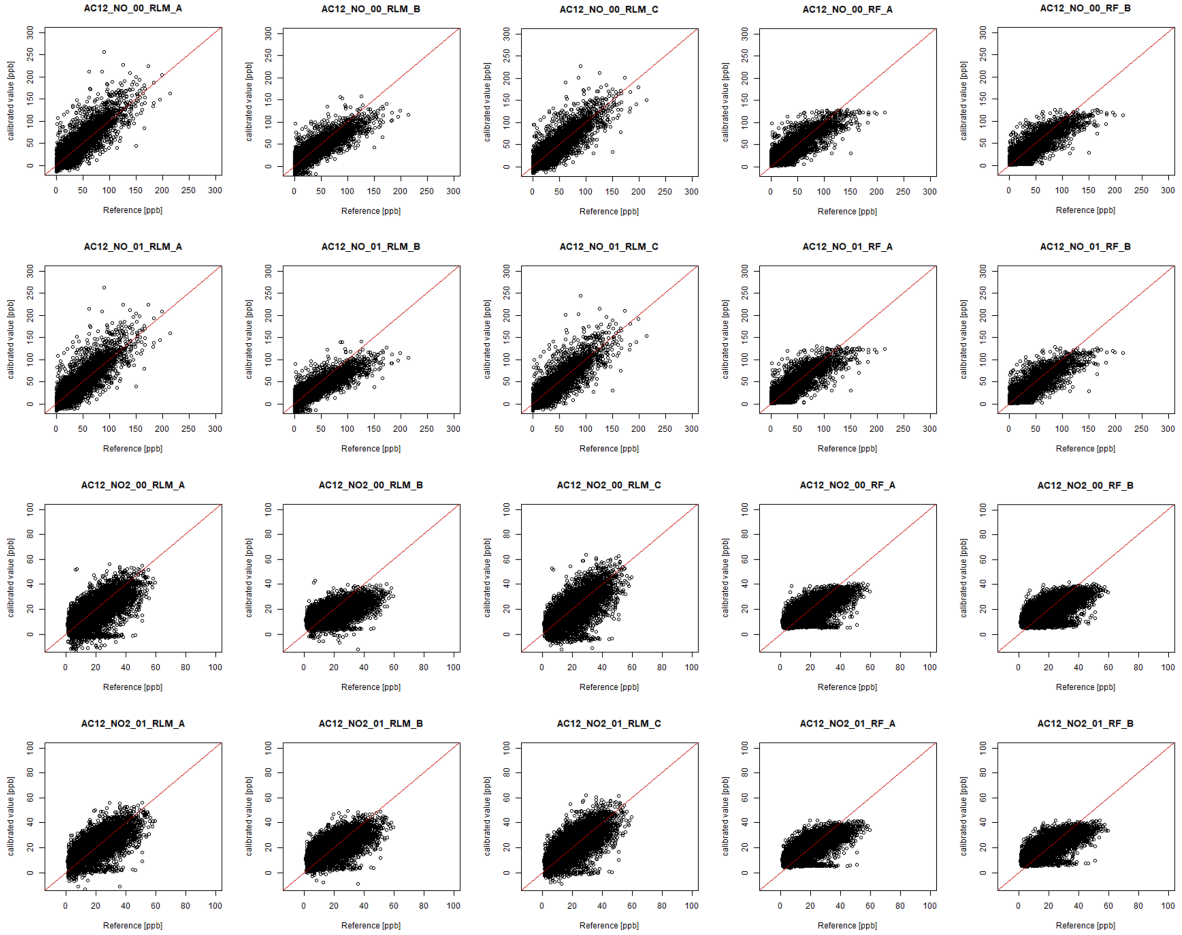


Figure 26: Correlation plots of AC012 for the measured concentration in the calibration site during 12.12.2019 - 31.03.20 with reference concentration data from the Harkingen monitoring station.

E.3 Comparison to the reference data from continuous monitoring stations in Zurich

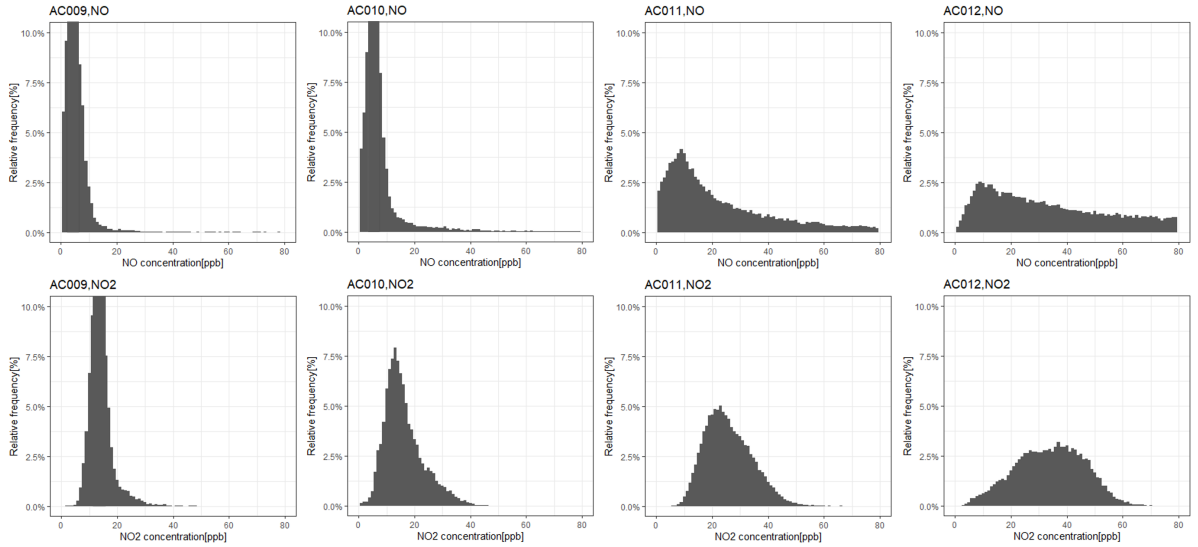


Figure 27: Histogram of NO and NO_2 concentration from the aircube sensor in the deployment sites during 14.12.18 - 09.08.19. The y-axis are limited to 10% and x-axis to 80 ppb for the comparison.

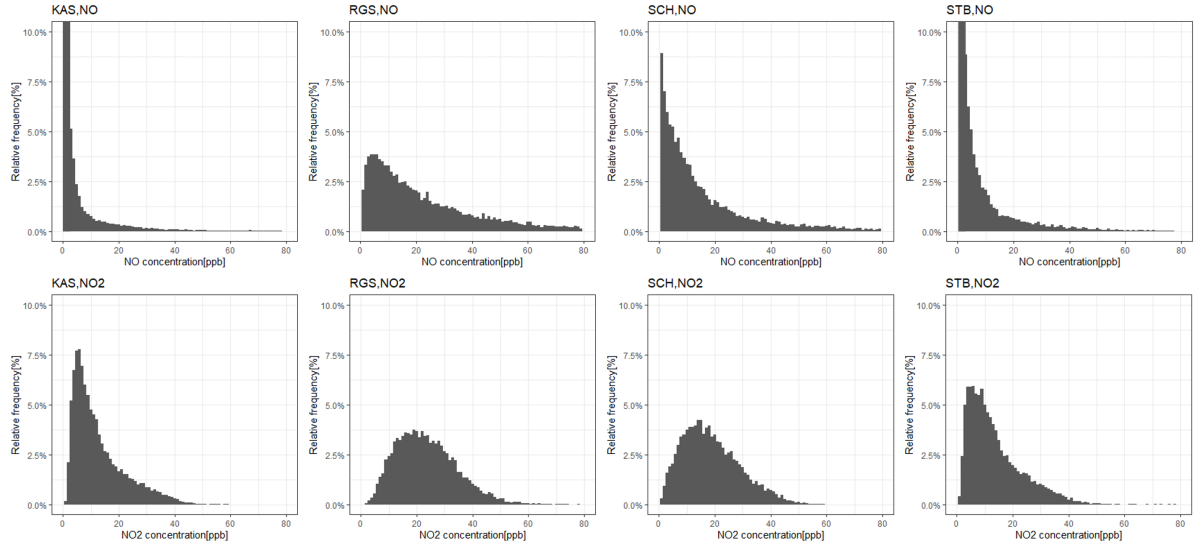


Figure 28: Histogram of NO and NO_2 concentration from the monitoring stations in Zurich (near Limmat and city center) during 14.12.18 - 09.08.19. The y-axis are limited to 10% and x-axis to 80 ppb for the comparison.

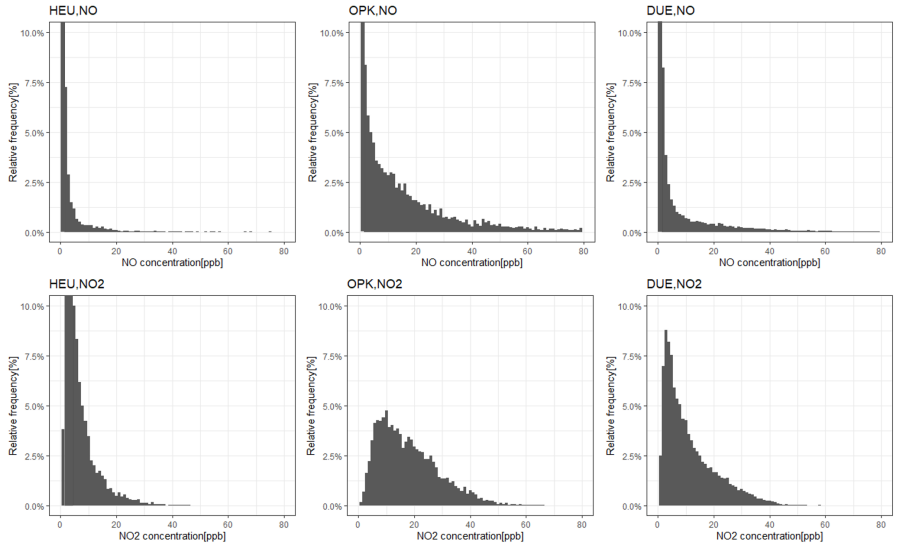


Figure 29: Histogram of NO and NO_2 concentration from the monitoring stations in Zurich (outer area) during 14.12.18 - 09.08.19. The y-axis are limited to 10% and x-axis to 80 ppb for the comparison.