Using Multi-parameters for Calibration of Low-cost Sensors in Urban Environment

Xinwei Fang
Department of Computer Science
University of York
xinwei.fang@york.ac.uk

Iain Bate
Department of Computer Science
University of York
iain.bate@york.ac.uk

Abstract

Measurements that were taken at appropriate spatial and temporal resolution are important for understanding urban environment. However, due to cost issues, most of current monitoring sensors are sparsely deployed and are not able to provide sufficient spatial resolution. As an alternative solution, low-cost sensors that cost several orders less than the current sensors have been exploited, providing much higher spatial resolution with relatively low cost. However, the data from low-cost sensors are widely reported to be deficient, resulting in the calibration of low-cost sensors being more difficult. In this work, key challenges of calibration of low-cost sensors were identified, and the limitations of current calibration methods were discussed. A multi-parameter calibration that not only utilises cross-sensitive parameters but also considers other relevant parameters was proposed. The stepwise regression method with interaction term was then proposed to systematically select optimal parameters for the calibration. The evaluations that were carried out in both city centre and outside of city centre have shown a great advantage of using the proposed method. It shows significantly better result than the existing methods, in terms of improved root mean square errors and better linearity between the calibrated trace and the reference.

Keywords

Low-cost sensors, multi-parameter calibration, urban environment

1 Introduction

Air quality deterioration has become a great concern in modern society. An increasing number of health problems, such as respiratory and cardiovascular diseases, are believed to be associated with prolonged exposure to hazardous air [40, 3]. This issue is especially relevant to urban residence as more than 50% of the world's population live in cities and the majority of cities are heavily polluted [45]. In order to maintain and improve human welfare and well-being, limitations of major pollutants are clearly stated in the newly introduced regulations and guidelines [9]. The regulatory limitations serve as a reference and govern the maximal pollution levels in a city. Therefore, it is essential to have measurements for the urban environment, and such measurements are currently provided by urban environmental monitoring networks [7, 33].

Urban environmental monitoring networks consist of a number of high quality sensors which are normally managed and maintained by government authorities [7, 33, 39]. The use of high quality sensors maximises the quality of monitored data in terms of precision and accuracy. However these sensors are expensive and require regular manual calibrations to ensure sustained sensor performance [9]. Thus, only a limited number of sensors can be afforded and deployed in a monitoring network. A further issue is urban environments are dynamic and unpredictable as it is a result of both emission and transmission of pollutants [20]. In urban environments pollution concentrations can vary significantly over small spatial and temporal scale due to the presence of dynamic pollution sources (cars), irregularity of urban topology and harsh physical conditions [30]. The complexity of the environment combined with having few high quality sensors results in high uncertainty of result in obtaining the urban pollution level [30]. Therefore, measurements that were taken at appropriate spatial and temporal resolution are essential.

As a result, dense networks constructed by using low-cost sensors, have been suggested and implemented in many studies [16, 15, 23]. It has been shown they can provide enhanced accuracy and confidence over current practices [29, 13]. However, it has been widely reported that the collected data from low-cost sensors suffers from large data uncertainties, relating to low data precision and accuracy, high percentage of data outliers and low correlation to reference [5, 27, 21]. It suggests that low-cost sensors may require comprehensive calibration processes in order to compensate for those reported issues and to ensure the data accuracy.

Calibrations of low-cost sensors have been intensively investigated in the past decade. Considering low-cost sensors are densely and broadly deployed, using conventional manual calibration methods is not practical and appropriate. Hence, automatic and semi-automatic calibration methods

International Conference on Embedded Wireless
Systems and Networks (EWSN) 2017
20–22 February, Uppsala, Sweden

ISBN: 978-0-9949886-1-4

have been widely used [2, 4, 32]. Sensor calibration is a process to determine a mathematical function, a calibration function, that describes the relationship between independent variables (uncalibrated parameters) and the dependent variable (reference). In most applications of sensor calibration, the calibration function considers and uses only the parameter of interest. The function is determined by directly comparing an uncalibrated data trace (the parameter of interest) to its reference. However the correlation between these can be weak [18, 19], leading to unsatisfactory accuracy [18]. As an alternative, a calibration method that utilises information from the cross-sensitive parameters have been proposed, and results have shown a significant improvement of calibration result [31, 36, 35, 41]. Cross-sensitive parameters are defined as the parameters that are cross-sensitive to the parameter of interest. However, to the best of our knowledge, the automatic determining of cross-sensitive parameters for a calibration process is not suggested in previous research, resulting in high possibility of including inappropriate parameters in the calibration. As using inappropriate parameters in a calibration process could result reduced accuracy and biased result of calibration [22, 10], it is an important issue to be addressed.

Furthermore, parameters, which are not cross-sensitive to the parameter of interest but are related to them, such as noise level, were not considered in the previous works. We believe that considering such parameters could provide key information on calibration of sensors. For example, noise level could related to the volume of traffic, and thus can be used for helping calibration of NO_2 as they are mainly generated by cars. Thus, we expect considering such parameters in the calibration process the calibration result could be further improved.

Parameters, like *noise*, that have unknown interaction with the parameter of interest but could provide important information for calibration of sensors, are called relevant parameters in this work. Multi-parameters in this work is defined as parameters that belong to either cross-sensitive parameters or relevant parameters. Multi-parameter calibration is then defined as a calibration process that utilise multi-parameters.

Thus the contributions of this work are: 1) to identify the main challenges in calibration of low-cost sensors; 2) to show the importance of using multi-parameter in calibrating low-cost sensors; 3) to show that calibration results are influenced by using different quality of data; and 4) to propose a method for calibration of sensors that systematically utilises multi-parameters and obtain an improved result than state of the art approaches.

This paper is organised as follows. In Section 2, data features of low-cost sensors are analysed and challenges of calibration of low-cost sensors are identified. The state-of-the-art works in calibration of low-cost sensors are reviewed in Section 3 and practical limitations of existing methods are also summarised. In Section 4, an automatic calibration method is proposed and it is evaluated and analysed in Section 5. Finally, conclusions and further suggestions are stated in Section 6

2 Understanding the Nature of Data

Data features depend on many factors and variation of data features can influence the result of sensor calibration significantly. In this section, key data features that directly and indirectly affect calibration of sensors are investigated, and how these data features vary by sensors and locations are also illustrated.

The use of sensors and locations of measurements in this work are firstly introduced in Section 2.1. The variation of data by using different sensors and locations is then illustrated in the Section 2.2. In Section 2.3 and Section 2.4, issues that potentially affect the result of multi-parameter calibration are explained. Finally, issues of calibration of low-cost sensors are summarised in Section 2.5.

2.1 Use of Sensors and Location of Measurements

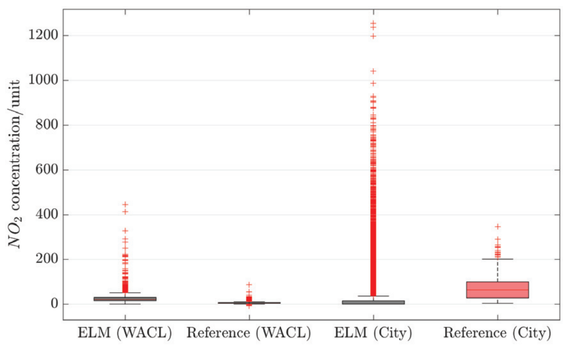
In this work, two types of sensors, high quality sensors and low-cost sensors, are placed at two locations, which are differentiated by environmental complexity. High quality sensors, which are also considered as references, are managed by the Wolfson Atmospheric Chemistry Laboratory (WACL), University of York, and Department for Environment, Food and Rural Affairs (Defra). They were calibrated before the experiment to ensure the precision and accuracy of data. For the low-cost sensors, ELM sensors, a product from Perkin Elmer, are used [26]. Both types of sensors are able to monitor multiple parameters including NO_2 and O_3, which allows multi-parameter calibration to be performed. To support cross comparison between the two types of sensors, the output of both are converted to concentration per unit.

Two locations were selected according to the environmental complexity. One location was selected on the top of the WACL building, which is on the university campus outside the city centre, in a relatively mild area. Another location was in the centre of the city of York (Fishergate), next to one of the busiest roads, where the environmental conditions are highly dynamic. These two locations will be referred to as WACL and City respectively in the following context. At both locations, high quality sensors were co-located with low-cost sensors within a meter range to minimise the spatial influences.

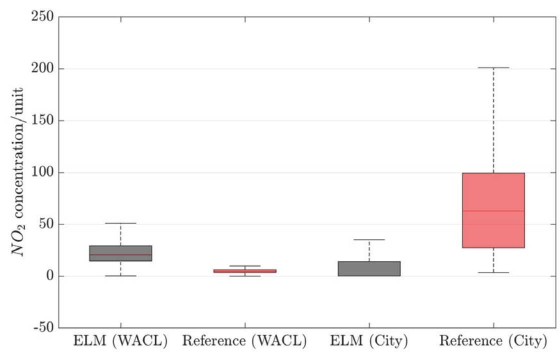
2.2 How and Why Data Varies

The type of sensors and the variation of environments are identified to be most influential factors for data variation [27, 21]. In this section, standard boxplots of NO_2 data are used to illustrate how data varies across different types of sensors and locations. The results are shown in Figure 1.

Figure 1a is plotted by using raw dataset. Extreme values, which are also referred to as outliers in this work, are represented by crosses in Figure 1a. The variation and scale of outliers can be clearly seen across sensors and locations. Outliers in the environmental data can be caused by many factors, such as interference from the environment or the malfunctioning of sensors. However, data from low-cost sensors contains more outliers in general, and the variability of those outliers is much larger; Comparing the same type of sensor at different locations, a larger variation of outliers can be observed in the City. Such systematic patterns of outliers between sensors and locations in Figure 1a indicate that difference of outliers in each sensor are related to the type of sensors and location of measurements. The result suggests that higher magnitudes and variations of outliers would easily occur in a sensor located in more complex environment than less complex environment, and occur in low-cost sensors than high quality sensors in general.



(a) Raw data



(b) Data without extreme values

Figure 1: Standard box-plots for data traces of NO_2 from four sensors at two locations

Figure 1b was rescaled from Figure 1a and plotted when extreme values were excluded. Thus, the variation of data is easier to see. Comparing sensors in the WACL in Figure 1b, both boxplots show equally spaced percentile values. Data from the reference sensor shows much less variation than the low-cost sensor, which confirms that measurements from reference sensors are more precise than the low-cost sensors. The difference in environmental conditions between WACL and City can be illustrated by comparing reference sensors at two locations in Figure 1b. It shows the variability of data in City is much higher than WACL and suggests the environmental interference from City could be higher than WACL. The result is also an evidence to explain why much higher magnitude of outliers are generally presented in City than WACL. Comparing sensors in the City, it is noticed that the data from low-cost sensors are highly skewed. ELM(City) boxplot in Figure 1b shows that the lower whisker, first quantile and median values of the boxplot are all at zero, which suggests that around 50% of data from low-cost sensor in City are zero values. A large percentage of constant value in the datasets suggests that the data from low-cost sensors in City performed abnormally.

According to the previous study [28], the differences in measurements between different types of sensors are caused by the sensitivity and selectivity of sensors, which are also believed to be associated with the use of sensing materials, sensor designs and environmental conditions. Sensitivity of a sensor is defined as the ability of a sensor to sense small changes of pollution concentration. Low sensitivity of a sensor could result in a low level of pollutants are incorrectly measured and further result in a large percentage of constant values in the data, like ELM(City) boxplot in Figure 1. The selectivity of a sensor is defined as the ability of a sensor to differentiate different pollutants. Low selectivity of a sensor can result in measurements are influenced by other substances in the air and are cross-sensitive to other parameters. Furthermore, environmental interference can have a large influence on both sensitivity and selectivity of sensors, which results in variation of data from City being more significant.

As low-cost sensors generally have lower sensitivity and selectivity, cross-sensitive parameters are more likely to occur in low-cost sensors. As a result, their data are more influenced by environment and are often less satisfactory than high quality sensors, especially in a complex environment, like a city centre [18, 28, 6, 11].

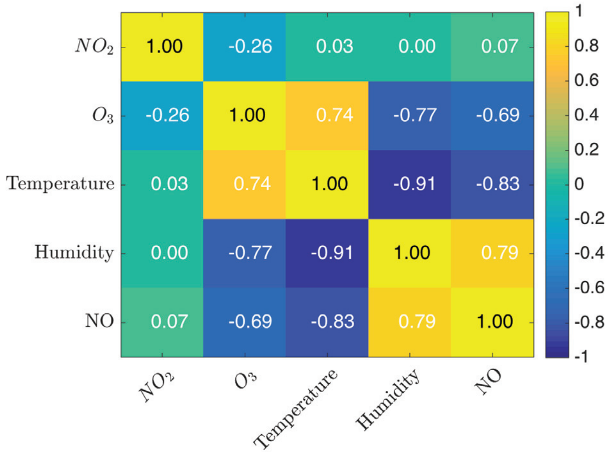
2.3 In-Situ Calibration

Cross-sensitive parameters of a sensor reduce the precision of measurements, and further affect calibration results that are obtained using only a single parameter of interest [18]. Thus, it is necessary to consider cross-sensitive parameters in the calibration of low-cost sensors [36, 35, 18]. However, unpredicted relevance among cross-sensitive parameters could result in utilising of cross-sensitive parameters in a calibration difficult.

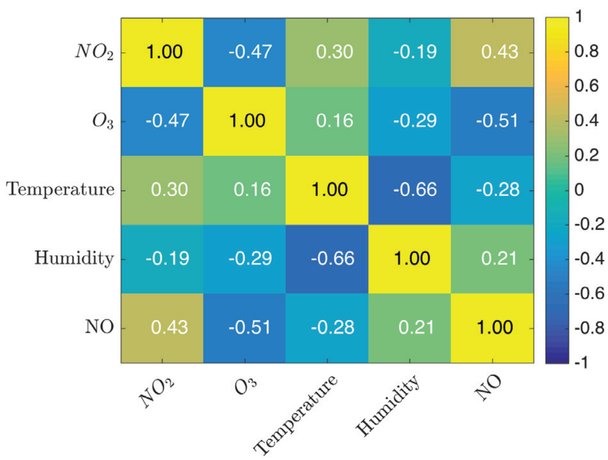
Figure 2 illustrates the cross-correlations of monitored parameters from an ELM sensor at two locations, WACL and City. The numbers showing on each figure indicate the correlation coefficients of two parameters which are indexed by rows and columns, and the colour intensity indicates the strength of correlation. The correlation coefficient represents a quantitative measure of the linear dependence between two parameters. Thus, the change of correlation can be considered as change of relevance between parameters. Comparing Figure 2a with Figure 2b, it can be noticed that some correlations change significantly when the location varies, such as NO and NO_2, and it indicates that the relevance between parameters is also different. It implies that a calibration method that works in one environment may need to use different parameters or coefficients for the calibration functions in other environments. Thus, the result indicates the importance of conducting a calibration in-situ.

2.4 Using Multi-parameters in Calibration

Multi-parameters include both cross-sensitive parameters and relevant parameters. The use of multi-parameters depends on availability of the parameters and relevance be-



(a) ELM sensor from WACL (WACL)



(b) ELM sensor from Fishergate (City)

Figure 2: Cross-correlation of parameters

tween the parameters. The availability of parameters is determined by the use of sensors. As different sensors measure different parameters, it may not be feasible to use exact parameters that used in the previous work due to the availability of parameters. The relevance between two parameters means the variation a parameter is related to another parameter. It indicates that the calibration result could be improved accordingly by including such parameters. As the cross-sensitive among parameters vary by locations, like shown in Figure 2, it suggests that the relevance among parameters is also different. Thus, some parameters that determined in the previous studies may not be appropriate parameters for this work. Furthermore, low-cost sensors could behave abnormally in complex environments, like ELM(City) boxplot in Figure 1b. Using data from such sensors could result inappropriate parameters are included in the calibration. According to the previous studies [37, 22, 10], adding inappropriate parameters in the calibration process would increase uncertainty of coefficients for the calibration functions and further bias the result of calibration. As it has high chance that in real practice inappropriate parameters are included in multi-parameter calibration, it is necessary to perform a systematic parameter selection for multi-parameter calibration.

Another issue is inter-relationship among multi-parameters. In the previous work [18], all cross-sensitive parameters are assumed to be independent. However, in real practice, parameters are unlikely to be fully independent. Figure 2 shows that only humidity and NO_2 in WACL are fully independent variables as correlation coefficient is zero. In other circumstance, the variation of one parameter may be dependent on other parameters. Using level of noise and concentration of NO_2 as an example, the level of noise could indicate the volume of traffic, which implies that the higher the noise level could represent the higher volume of traffic. As the NO_2 concentration is closely linked to the volume of traffic, the concentration of NO_2 could vary differently when the level of noise is different [12]. Therefore, it is necessary to consider the relationship among multi-parameters in order to obtain an optimal result.

2.5 Summary

The use of the sensors and the selection of locations are identified to be the most influential factors for data features. Data features in terms of magnitudes of outliers and variation of data can be significantly different when those two factors change. Furthermore, the cross-sensitivity among parameters can be differently at different locations. Therefore, it is necessary to perform calibration process for new sensors in-situ.

Existing studies have shown the importance of using cross-sensitive parameters. However, it is likely to include inappropriate parameters in the process and bias the result of calibration, as explained in Section 2.4. It raises the necessity of having a systematic method for determining multi-parameters. Moreover, considering the inter-relationship among multi-parameters is believed to maximise the result of multi-parameter calibrations.

3 Related work and existing limitations

Calibration of low-cost sensors, in general, can be classified as micro-calibration and macro-calibration [38, 43].

Macro-calibration does not require a reference as it only utilises the consistency of the nearby environment and maximizes the similarity of measurements among the co-located sensors [2]. This method is flexible and able to calibrate large and dense networks with relatively low cost. However, the calibrated result from macro-calibration is only relative to the consistent measurements among sensors, which may be biased if all sensors systematically drift. Furthermore, considering the data variation and data quality of low-cost sensors in Figure 1a, the required consistency of measurement may not be obtainable, which suggests that macro-calibrations is not suitable for calibrating low-cost sensors, like ELM, or in a dynamic environment, like a city centre.

Micro-calibration, on the other hand, requires the presence of a reference and does not assume sensor behaviours like those in macro-calibration. Therefore, the obtained calibration function is more accurate and the calibrated result is more close to the expected value [27, 32]. Therefore, micro-calibration is used to calibrate low-cost sensors in this work.

Traditional calibration considers and uses only the parameter of interest. The function is determined by directly comparing an uncalibrated data trace to its reference. However, data features that are mainly caused by cross-sensitive parameters, as explained in Section 2.2, make calibration of low-cost sensors using only the parameter of interest insufficient in terms of calibration accuracy [34]. Thus, it raises the importance of using cross-sensitive parameters in the calibrations.

Studies on cross-sensitive of parameters have been mainly performed in an controlled environment. Morsi discovered that changing temperature and concentration of gases influences the selectivity of gas sensors, such as CH_4 and CO_2 sensor [24]. Furthermore, Martin noticed that selectivity and sensitivity of NO_2 and CO sensors have a dependence on the temperature [19] and Losch discovered that a large cross sensitivity exists in between O_3 and temperature [17]. The studies show environmental parameters are broadly cross-sensitive to each other, and their results indicate the importance of including cross-sensitive parameters in calibration process.

The use of relevant parameters, such as noise or traffic data have been widely used in environmental modelling [14, 1, 42]. However, to the best of our knowledge, such information is rarely considered in the calibration of low-cost sensors.

Eugster [8] utilised cross-sensitive parameters and eliminated the effects from humidity and temperature by using the method in [24]. The result shows slightly improved calibration result of sensors. Spinelle tested different sensors, mainly O_3 and NO_2, and performed calibration in an open environment by using cross-sensitive parameters that were provided by high quality sensors [36, 35]. Spinelle utilised joint information between O_3 and NO_2 in their calibration and the results indicate the calibration result of O_3 was significant improved where the result of NO_2 was only improved slight. In work [18], Maag also performed cross-sensitive calibration on several different low-cost sensors in an open environment. They discovered that parameters, like NO_2, O_3, temperature and humidity, are cross-sensitive to each other. They also identified CO is an irrelevant parameter for calibrating both NO_2 and O_3 sensor as adding the parameter of CO did not improve the calibration results. Their result further supports that there is high chance to include irrelevant parameter in real practice. Both Spinelle and Maag's works suggest that parameters generally have higher correlation with O_3 than NO_2 and it suggests that an NO_2 sensor is less related to other parameters and could be more difficult to calibrate [36, 35, 18].

Both Spinelle and Maag suggest considering cross-sensitive parameters in the calibration of pre-deployment of sensors [34, 18]. However, as summarised in Section 2.5, a calibration function that works in one location may not be applicable for another location. Hence, it is necessary to test the calibration method in a working conditions. Their works did not suggest an automatic method to determine the cross-sensitive parameters from available dataset. Thus, their work can not be directly applied in a general application. Furthermore, to the best of our knowledge, the existing works did not considered inter-relationship among cross-sensitive parameters, which suggests that calibration result can still be improved. A further issue is that the data from cross-sensitive parameters are currently provided by high quality sensors. In real practice, those high quality sensors are not always accessible. Therefore, it is necessary to understand how using data from low-cost sensors and high quality sensors would affect multi-parameter calibrations.

4 Method

Utilising cross-sensitive parameters in calibration of low-cost sensors have shown a great advantage than the calibrations that use only a single parameter [34, 18, 35]. In this work, a regression based multi-parameter method will be investigated as it has been suggested to have a better performance in practice than other methods [18].

In this section, inter-relationships among multi-parameters are firstly addressed in Section 4.1. A two-way interaction term from any two parameters is introduced, which has been used in [12]. Then, stepwise regression is proposed in Section 4.2 which avoids including inappropriate parameters by systematically adding parameters into the regression. As a result, the proposed method are expected to automatically utilise parameters (cross-sensitive parameters and relevant parameters) from any available dataset and produce better calibration result than the state of art approaches that introduced in [18].

4.1 Relationship among Parameters

Section 2.4 has shown that the variation of NO_2 can be different at different values of noise. Therefore, such relationship needs to be considered in the calibration.

Assume a dependent variable, Y, which can be the reference data of NO_2, and independent variables, X_1 and X_2, which can be *noise* and uncalibrated NO_2. Using the method from [18], the calibration function with corresponding coefficients β can be constructed as Equation 1:

$$Y \sim \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 \quad (1)$$

From the equation, both X_1 and X_2 are independent, which means the change of X_1 would not affect X_2. However, in reality that may not be a case. One possibility is that at higher noise level, the variation of uncalibrated NO_2 tend to be more dramatic, whereas at lower noise level, the variation of uncalibrated NO_2 is much less, considering the noise level is linked to the volume of traffic. It suggests that the variation of uncalibrated NO_2 is also depend on level of *noise*.

In order to solve the issue, a two-way interaction term is introduced [12]. Adding interaction terms into the calibration function, the new equation will look like Equation 2:

$$Y \sim \beta'_0 + \beta'_1 \cdot X_1 + \beta'_2 \cdot X_2 + \beta'_3 \cdot (X_1 \times X_2) \quad (2)$$

and the calibration function can be re-write as Equation 3:

$$Y \sim \beta'_0 + \beta'_1 \cdot X_1 + (\beta'_2 + \beta'_3 \cdot X_1) \cdot X_2 \quad (3)$$

In this case, the slopes of the equation between reference NO_2 (Y) and uncalibrated NO_2 (X_2) are different for the different level of noise (X_1). Hence, the variation of NO_2 is related to the level of noise and β'_3 indicates how different those slopes are. Therefore, adding interaction terms is believed to be able to maximise inter-relationship among multi-parameters.

4.2 Stepwise Regression

Least square regression is a method that has been widely applied to determine sensor calibration functions [32] and it has been suggested to work better for calibration of using cross-sensitive parameters than other methods [18]. Least square regression is worked as the best approximation in terms of least square errors between a dependent variable and prediction result.

The least square method can also be used on multiple independent variables, which is called Multiple Least Square (MLS) regression. As more than one independent variables are used, the approximation is conducted in a multi-dimensional space as Equation 4, where Y stands for a dependent variable, which can be considered as a reference of a uncalibrated parameter, and X stands for independent variables, which can be considered as uncalibrated parameter and other available terms (measured parameters and interaction terms). β stands for corresponding coefficients in the function where n represents the index for the number of independent variables.

$$Y \sim \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_n \cdot x_n \tag{4}$$

MLS uses functions in Equation 5 and 6 to determine optimal coefficients, where i indicates a time instance and $i \in 1, 2, ..., N$ and N equals to the number of time stamps in data.

$$r_i = Y_i - (\beta_{0i} + \beta_{1i} \cdot X_{1i} + \beta_{2i} \cdot X_{2i} + \cdots + \beta_{ni} \cdot X_{ni}) \tag{5}$$

$$minimize \sum_{i=1}^{N} r_i^2 \tag{6}$$

Equation 4 utilises all available variables without performing a systematic analysis to determine optimal parameters, it could easily include inappropriate parameters and bias result of calibration. As a result, stepwise regression is proposed to automatically utilise multi-parameter from any available dataset, which enables the method to be applied directly to most of applications.

Stepwise regression is a regression method that is similar to the MLS regression. The key difference is instead of using all parameters from dataset, stepwise regression performs a stepwise process to select suitable parameters from the dataset. The method starts with only one independent variable and at each step a new variable from the dataset is added into the regression. The result of regression is compared between the current step (with the new variable) and the previous step (without the new parameter). If a newly added variable has a significant positive contribution to the process in terms of calibration accuracy, this parameter is included in the process. If the result fails to show a significant improvement, then this newly added parameter is excluded from the subsequent process. Sum of Squared-Error (SSE) and F-test is used to determine the relevance of the newly added parameter, which was also used in [18].

The stepwise regression between any two steps is performed according to Equation 7 - 9:

$$SSE' = Y - \beta'_0 + \beta'_1 \cdot x_n \tag{7}$$

$$SSE'' = Y - \beta''_0 + \beta''_1 \cdot x_n + \beta''_2 \cdot x_{n+1} \tag{8}$$

SSEs between steps are calculated using Equations 7 and 8 and statistical significance between two steps are determined by using F-test. The SSE is used to indicate the result of the regression, where larger SSE value represents for less precision of approximation. Statistical significance (p $<$0.05) is used in this case.

The null-hypothesis of each comparison is that adding the next variable into the regression will not make a significant difference to regression result. The alternative hypothesis is that adding the next variable will make a significant difference. If tests fail to accept the null-hypothesis, then the alternative hypothesis is accepted. Furthermore, if a positive contribution is determined, the new parameter is then added to the calibration function, the logic is indicated in Equation 9

$$Y \sim \begin{cases} \beta''_0 + \beta''_1 \cdot x_n + \beta''_2 \cdot x_{n+1}, & \text{if } SSE'' < SSE' \ \& \ p < 0.05 \\ \beta'_0 + \beta'_1 \cdot x_n, & \text{otherwise} \end{cases} \tag{9}$$

This process continues until all available parameters are tested, which also includes interaction terms that introduced in Section 4.1. Thus, the proposed method could minimise the uncertainty of calibration function by excluding inappropriate parameters from the process. Another advantage of using stepwise regression is that it does not require pre-determined knowledge of input data, like number of available parameters. Hence, it could be applied in general cases where the use of sensors, the number of measured parameters and locations are different. The use of method is illustrated in Algorithm 1.

5 Evaluation

In this section, the methods are evaluated. The experiment set-up and the use of data is first introduced in the Section 5.1. The results of the proposed method are compared against to the state of the art approach used in [18] in WACL. Then, effect of calibration result relating to data qualities of multi-parameters of is tested in Section 5.3. Finally, the performance of the proposed method in real environment is evaluated in the city centre.

In this work, calibration of NO_2 is considered as it is widely reported to be more difficult to calibrate than other parameters, and it is one of major pollutant in urban environment [36, 35, 18].

5.1 Dataset and Pilot Experiment

The type of sensors and selection of locations of this experiment is based on Section 2.1. Reference sensors at both location measure NO_2, NO and O_3. Two versions of ELM

Data: Reference: $Y_{m\times1}$ Uncalibrated data trace: $x_{1_{m\times1}}$ Available parameters : $x_{2_{m\times1}}$ to $x_{n_{m\times1}}$. m is for numbers of samples in each variable and n is for total number of available parameters

for *i = 1 to n* **do**
  for *j = i+1 to n* **do**
    $terms_{i,j} = x_{j_{m\times1}}\ x_{i_{m\times1}}$
  end
end

Result: Obtaining two-way interaction terms for all parameters, $terms_{m\times t}$, t is for number of interaction terms

$X_{m\times(t+n)} = terms_{m\times t} + x_{m\times n}$;

Result: Combining interaction terms and measured parameters as independent variables, X

for *k = 1 to numbers of ((t+n)-1)* **do**
  SSE(k) = Y - MLS($X_{1_{m\times1}}$ to $X_{k_{m\times1}}$);
  SSE(k+1) = Y - MLS($X_{1_{m\times1}}$ to $X_{k+1_{m\times1}}$);
  p = significance(SSE(k),SSE(k+1))
  if *SSE(k+1) < SSE(k) & p < 0.05* **then**
    $X_{k+1_{m\times1}}$ is used in multi-parameter calibration
  else
    $X_{k+1_{m\times1}}$ is NOT used in multi-parameter calibration
  end
end

Result: Using stepwise regression to determine parameters and calibration function

algorithm 1: A pseudo code for the method

are used at different locations and they measure different parameters. Thus, it can be used to test the performance of proposed method across sensors and locations. ELM sensor at WACL measures parameters of NO_2, O_3, *Temperature*, *Humidity*, *dust*, *VOC* and *noise*. ELM sensors in the City measures parameters of NO_2, NO, O_3, *Temperature* and *Humidity*.

Calibration function with linear model is commonly used in most of calibration methods [2, 32]. However, other works suggest that a non-linear model could be more optimal than a linear model in their works [11, 25, 46]. Thus, it is necessary to test the use of model before the experiment. In the pilot experiment, various models from linear up to higher order polynomial were tested. The linear model shows a great advantage over other models in our work, as the linear model not only provides good calibration accuracy but also minimises computational time.

5.2 Evaluation in WACL

A number of co-located ELM sensors were previously deployed on WACL, as illustrated in Figure 3. Due to the availability, eleven sensors are used to test the variation of calibration results across sensors. Two months worth of data is used, and in order to minimise the effect from natural variation of the data, cross-validation is used to obtain the results.

Figure 3: ELM sensors on top of WACL building

Multiple rounds of cross-validation are performed using different partitions of data, and the calibration result of a sensor, in terms of one RMSE value in each boxplot in Figure 4, is averaged over those rounds.

Root-mean-squared error (RMSE) between calibrated trace and reference is used to determine the performance of calibration. The smaller the RMSE value means calibration accuracy is better. In this work, every one unit of RMSE improvement represents about 20 percentage better accuracy of calibration. The calibration results in WACL is shown in Figure 4. Each boxplot presents variation of calibration result over eleven sensors, the results obtained by different methods are differentiated by colours. Seven groups of results indicate seven datasets with different parameters are used. The seven parameters are determined by the number of measured parameters from ELM sensor. Datasets are constructed by gradually adding a new parameter into the previous dataset, one parameter at each time. The first group only NO_2 is considered, the second group, O_3 is added into the previous group, and so on. Therefore, in the last group, total seven parameters are used. The sequence of adding a new parameter is random, which is the order of $O_3, H, T, dust, VOC$ and *noise*.

In Figure 4, the importance of using multi-parameter calibration can be observed by comparing the first group with the rest of groups. The significant improvement of calibration result is obtained when more than one parameter is used in the calibration.

By gradually adding a parameter into the calibration process, the MLS method shows a reduced accuracy for some sensors when T is added into the dataset. It confirms that adding inappropriate parameters can indeed reduce the calibration result. On the contrary, as the proposed method systematically utilises parameters, the calibration accuracy are gradually improved.

Parameters of O_3, H, and T had been identified as the cross-sensitive parameter for NO_2 in the previous studies [18]. In the Figure 4, result confirms that parameters of O_3, H, and T are important parameters for calibrating NO_2 as they have positive contribution to the calibration in this study as well. However, parameters of *dust*, *VOC* and *noise*, which are considered as relevant parameter in this work, were not used in the previous studies. Figure 4 shows the result of calibration does not change when these parameters are added in the MLS method, but the result of the pro-

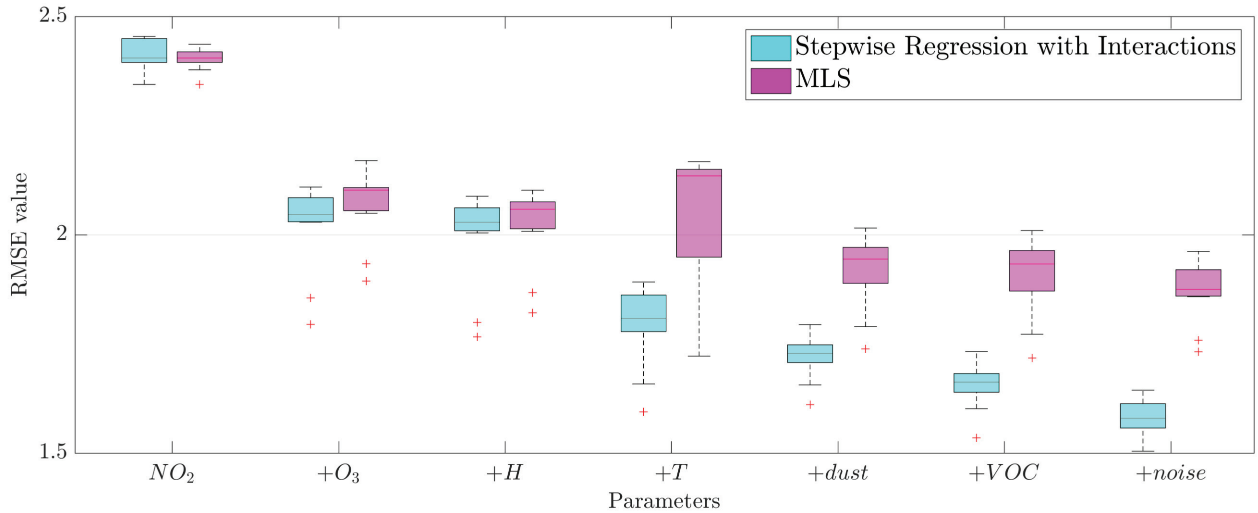


Figure 4: Calibration of a sensor in WACL

posed method is gradually improved when these parameters are considered. The difference in results between two methods confirms the our previous assumption that utilising inter-relationship among parameters can indeed maximise calibration results.

Figure 4 indicates the importance of using multi-parameters in calibrations and suggests the proposed method generally have better performance than MLS method, especially when the relevant parameters are introduced.

5.3 Influences from Data Quality

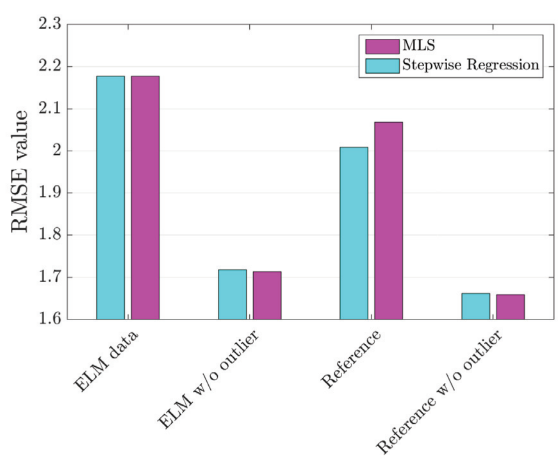


Figure 5: reference Ozone Vs ELM Ozone

As mentioned in Section 3, the previous works rely on high quality sensor to provide data for cross-sensitive parameters. However, high quality sensors are not always accessible in real practice to provide such information. Therefore, using multi-parameters provided by low-cost sensors seems to be unavoidable. In order to understand how quality of data from multi-parameters would affect the calibration result, this evaluation is performed.

Two datasets, a high quality dataset from reference and a lower quality dataset from ELM are used to provide data for multi-parameters. As NO_2 and O_3 are the only two common parameters that are available for both sensors, O_3 is used as the only multi-parameter to calibrate NO_2.

We assume that high quality dataset has much better data quality in terms of precision and accuracy than the dataset from low-cost sensor. Therefore, the effect on calibration result that caused by the quality of O_3 can be differentiated by comparing the calibration results that use O_3 from different type of sensors. Furthermore, we remove the extreme values from the O_3 datasets to evaluate how outliers from multi-parameter would affect the calibration of NO_2.

As data distributions of parameters are unclear, a non-parametric method was used to remove outliers. Outliers are considered as extreme data values that above or below 1.5 times of a distance between two quantiles of the boxplot [44]. Both stepwise regression with interaction terms and MLS calibration methods are applied to each dataset to see how results vary by methods and the result of calibration is shown in Figure 5.

In Figure 5, two colours of box-plot indicates the two calibration methods and four blocks represent four datasets. ELM data indicates the parameter of O_3 is used from ELM and Reference indicates the parameter of O_3 is used from high quality sensor. ELM w/o outlier and Reference w/o outliers stands for the outlier from parameter of O_3 are removed.

In Figure 5, the variation of results between two methods is not significant. It indicates that the difference in calibration result is caused by the use of data but not methods. Comparing block one with block three, using low quality data and high quality data, a convincing enhancement of calibration result can be concluded when a higher quality data is used for multi-parameter. However, outliers, which are considered as extreme values, have an even higher impact on the calibration accuracy. Comparing datasets of O_3 with and without outliers, the results show a clear pattern that datasets after outlier removal have better calibration accuracy than another ones.

Figure 5 suggests that precision and accuracy of O_3 do have influence on the calibration result but the outliers of O_3 have more impact on the calibration result. As removal of outlier in this section is for showing the calibration result without presence of extreme value, the quality of outlier removal is not evaluated. Thus, the outlier removal is not considered and used in the experiments of this work.

5.4 Calibration in City Centre

As mentioned in Section 2 and Section 3, it is necessary to calibrate sensors under the working conditions. Therefore, an ELM sensor which is calibrated at Fishergate, York. For this experiment, ELM sensors were placed closely to the reference station.

A month worth of data is used, where the first two weeks data are used for training and last two weeks data are used for testing. The proposed method is used and available parameters are NO_2, O_3, NO, $Temperature$ and $Humidity$ from ELM and NO_2 from reference. Figure 6 is a series of scatter plots between the ELM data and reference at Fishergate. The ELM data is expected to have a perfect linear dependence with the reference in the Figure 6 as they stand for the measurement of NO_2 at the same time and location. The correlation coefficient (R) between ELM data and reference is calculated and a line which determined by the least square method is also plotted. As the line in Figure 6 is calculated based on least square method, the equation of the line indicates the linearity between ELM data and reference, and the slope and the offset of the equation is expected to be close to one and zero respectively.

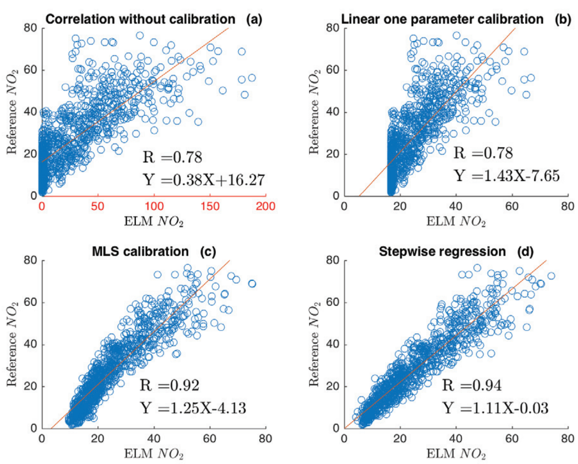


Figure 6: Calibration of a sensor in Fishergate

Figure 6a shows raw ELM with the reference. It is difficult to determine a proper fit between two readings and data range vary significantly in ELM sensor. The result also illustrates that the low-cost sensor in city centre is hard to calibrate.

Then, a linear calibration using only parameter of NO_2 is applied in the Figure 6b. As only a single parameter, NO_2, was used, the calibration result is hard to explain as there are large number of zero values at a certain point. It also indicates that using only a single parameter to calibrate low-cost sensor in urban environment is more likely to be insufficient.

Figure 6c shows the calibration result is significantly improved when MLS multi-parameter calibrations are considered and suggests the importance of using multi-parameter to calibrate low-cost sensors.

Figure 6d shows result when stepwise regression with interaction terms is used. The correlation between the reference and ELM is much stronger than using the MLS calibration. Furthermore, the linearity to the reference is also improved by comparing the slope and offset with Figure 6c, where the slope is much closer to one and offset is much close to zero. Therefore, it is able to conclude that the proposed method also works better than existing methods in urban environment.

6 Conclusions

The needs of high spatial resolution data encourages the use of low-cost sensors. Due to the sensor design and the use of low-cost material, the data quality from low-cost sensor is often reported insufficient in terms of data variation, outliers and precisions and accuracy. It makes calibration of low-cost sensors important but difficult, especially in urban environments.

In this work, the nature of data and how they influence the calibration of sensors are explained in Section 2. The environmental complexity, the use of sensors and data features, like high variation of data and outliers, are identified to be the main factors for calibration of low-cost sensors.

The evaluation at WACL is able to confirm the importance of using multi-parameters to calibrate low-cost sensors and suggests the importance of using interaction term to maximise the relationship among parameters. The result also supports our assumption that the result of calibration may not be optimal if includes inappropriate parameter or not utilises inter-relationship among multi-parameters. The proposed method shows a significant improvement of calibration result when relevant parameters are added, which suggests that relevant parameter should also be considered in the calibration. In the future work, more relevant parameters will be exploited, such as weather data.

Moreover, we identified that the outliers from data have more influence on calibration result than data precision and accuracy. This finding suggests the importance of outliers removal in the sensor calibration and it is essential to be addressed in the future work.

The calibration process that tested at Fishergate, York confirms that use of multiple parameters helps to obtain a better calibration result in the urban environment than using only a single parameter. The result also indicates the proposed method is able to work across locations and sensors, and work better than existing methods in general.

In contrast to existing work [18], our method can be directly applied on most of sensors and obtain better calibration results in various locations.

7 Acknowledgement

This work is funded by the European Unions Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 608014 (CAPACITIE)

8 References

[1] I. Aguilera, J. Sunyer, R. Fernández-Patier, G. Hoek, A. Aguirre-Alfaro, K. Meliefste, M. Bomboi-Mingarro, M.and Nieuwenhuijsen, D. Herce-Garraleta, and B. Brunekreef. Estimation of outdoor NO_x, NO_2, and BTEX exposure in a cohort of pregnant women using land use regression modeling. *Environmental Science & Technology*, 42(3):815–821, 2007.

[2] L. Balzano and R. Nowak. Blind calibration of sensor networks. In *Proceedings of the 6th International Conference on Information Processing in Sensor Networks*, pages 79–88. ACM, 2007.

[3] R. Brook, B. Franklin, W. Cascio, Y. Hong, G. Howard, M. Lipsett, R. Luepker, M. Mittleman, J. Samet, and S. Smith. Air pollution and cardiovascular disease: A statement for healthcare professionals from the expert panel on population and prevention science of the american heart association. *Circulation*, 109(21):2655–2671, 2004.

[4] V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak. A collaborative approach to in-place sensor calibration. In *Information Processing in Sensor Networks*, pages 301–316. Springer, 2003.

[5] Y. Cheng, X. Li, Z. Li, S. Jiang, Y. Li, J. Jia, and X. Jiang. Aircloud: A cloud-based air-quality monitoring system for everyone. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, pages 251–265. ACM, 2014.

[6] L. Clapp and M. Jenkin. Analysis of the relationship between ambient levels of O_3, NO_2 and NO as a function of NO_x in the UK. *Atmospheric Environment*, 35(36):6391–6405, 2001.

[7] EPA. Air quality. http://www.epa.ie/air/quality/#V̇mmzTxqLRPV, February 2015. [Online; accessed on 26-September-2016].

[8] W. Eugster and G. Kling. Performance of a low-cost methane sensor for ambient concentration measurements in preliminary studies. *Atmospheric Measurement Techniques*, 5(8):1925–1934, 2012.

[9] European Union. Directive 2008/50/ec of the european parliament and of the council of 21 May 2008 on ambient air quality and cleaner air for europe. *Official Journal of the European Union*, 2008.

[10] P. Filzmoser, M. Gschwandtner, and V. Todorov. Review of sparse methods in regression and classification with application to chemometrics. *Journal of Chemometrics*, 26(3-4):42–51, 2012.

[11] S. Han, H. Bian, Y. Feng, A. Liu, X. Li, F. Zeng, and X. Zhang. Analysis of the relationship between O_3, NO and NO_2 in Tianjin, China. *Aerosol Air Quality Research*, 11(2):128–139, 2011.

[12] A. Hayes. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press, 2013.

[13] R. Hijmans, S. Cameron, J. Parra, P. Jones, and A. Jarvis. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978, 2005.

[14] G. Hoek, R. Beelen, K. De, D. Vienneau, J. Gulliver, P. Fischer, and D. Briggs. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric environment*, 42(33):7561–7578, 2008.

[15] King's College London. Nowcast accuracy. http://www.londonair.org.uk/london/asp/NowcastConfidence.asp, February 2015. [Online; accessed on 26-September-2016].

[16] P. Kumar, L. Morawska, C. Martani, G. Biskos, M. Neophytou, S. Di, M. Bell, L. Norford, and R. Britter. The rise of low-cost sensing for managing air pollution in cities. *Environment International*, 75:199–205, 2015.

[17] M. Losch, M. Baumbach, and A. Schutze. Ozone detection in the ppb-range with improved stability and reduced cross sensitivity. *Sensors and Actuators B: Chemical*, 130(1):367–373, 2008.

[18] B. Maag, O. Saukh, D. Hasenfratz, and L. Thiele. Pre-deployment testing, augmentation and calibration of cross-sensitive sensors. pages 169–180. ACM, 2016.

[19] M. Martin, J. Santos, H. Vasquez, and J. Agapito. Study of the interferences of NO_2 and CO in solid state commercial sensors. *Sensors and Actuators B: Chemical*, 58(1):469–473, 1999.

[20] H. Mayer. Air pollution in cities. *Atmospheric Environment*, 33(24):4029–4037, 1999.

[21] M. Mead, O. Popoola, G. Stewart, P. Landshoff, M. Calleja, M. Hayes, J. Baldovi, M. McLeod, T. Hodgson, and J. Dicks. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment*, 70:186–203, 2013.

[22] T. Mehmood, K. Liland, L. Snipen, and S. Sæbø. A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 118:62–69, 2012.

[23] H. Messer, A. Zinevich, and P. Alpert. Environmental monitoring by wireless communication networks. *Science*, 312(5774):713–713, 2006.

[24] I. Morsi. A microcontroller based on multi sensors data fusion and artificial intelligent technique for gas identification. In *33rd Annual Conference of the IEEE, Industrial Electronics Society, IECON.*, pages 2203–2208. IEEE, 2007.

[25] S. Munir, H. Chen, and K. Ropkins. Modelling the impact of road traffic on ground level ozone concentration using a quantile regression approach. *Atmospheric Environment*, 60:283–291, 2012.

[26] Perkin Elmer. Elm sensor. https://elm.perkinelmer.com/, February 2015. [Online; accessed on 26-September-2016].

[27] R. Piedrahita, Y. Xiang, N. Masson, J. Ortega, A. Collier, Y. Jiang, K. Li, R. Dick, Q. Lv, and M. Hannigan. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques*, 7(10):3325–3336, 2014.

[28] C. Pijolat, C. Pupier, M. Sauvan, G. Tournier, and R. Lalauze. Gas detection for automotive pollution control. *Sensors and Actuators B: Chemical*, 59(2):195–202, 1999.

[29] R. Pope and J. Wu. Characterizing air pollution patterns on multiple time scales in urban areas: A landscape ecological approach. *Urban Ecosystems*, 17(3):855–874, 2014.

[30] B. Resch, M. Mittlboeck, F. Girardin, R. Britter, and C. Ratti. Live geography: Embedded sensing for standarised urban environmental monitoring. *International Journal on Advances in Systems and Measurements*, 2009.

[31] D. Reza, S. Nejad, M. Kabganian, and F. Barazandeh. Neural network calibration of a semiconductor metal oxide micro smell sensor. In *Symposium on Design Test Integration and Packaging of MEMS/MOEMS (DTIP)*, pages 154–157. IEEE, 2010.

[32] O. Saukh, D. Hasenfratz, and L. Thiele. Reducing multi-hop calibration errors in large-scale mobile sensor networks. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, pages 274–285. ACM, 2015.

[33] Seoul Metropolitan Government. Air quality information. http://english.seoul.go.kr/policy-information/environment-energy/air-quality-information/1-air-quality-information/, February 2015. [Online; accessed on 26-September-2016].

[34] L. Spinelle, M. Gerboles, and M. Aleixandre. Report of laboratory and in-situ validation of micro-sensor for monitoring ambient air-ozone micro-sensors, αSense, model:$B4$ O_3 sensors. *Publications Office of the European Union*, 26681, 2013.

[35] L. Spinelle, M. Gerboles, M. Villani, M. Aleixandre, and F. Bonavitacola. Calibration of a cluster of low-cost sensors for the measurement of air pollution in ambient air. In *SENSORS*, pages 21–24. IEEE, 2014.

[36] L. Spinelle, M. Gerboles, M. 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*, 215:249–257, 2015.

[37] A. Studenmund. *Using Econometrics: A Practical Guide*. Addison-Wesley Series in Economics. Pearson Addison Wesley, 2006.

[38] R. Tan, G. Xing, Z. Yuan, X. Liu, and J. Yao. System-level calibration for data fusion in wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 9(3):28, 2013.

[39] The Asian Institute for Energy, Environment and Sustainability. The air quality index for Seoul, South Korea. http://epi2012.yale.edu/indicators-in-practice/air-quality-index-seoul-south-korea, February 2015. [Online; accessed on 26-September-2016].

[40] P. Tolbert, J. Mulholland, D. Macintosh, F. Xu, D. Daniels, O. Devine, B. Carlin, M. Klein, A. Butler, and D. Nordenberg. Air quality and pediatric emergency room visits for asthma in Atlanta, Georgia. *American Journal of Epidemiology*, 151(8):798–810, 2000.

[41] S. Vaihinger and W. Göpel. Multi-component analysis in chemical sensing. *Sensors Set: A Comprehensive Survey*, pages 191–237, 2008.

[42] A. Wheeler, M. Smith-Doiron, X. Xu, N. Gilbert, and J. Brook. Intra-urban variability of air pollution in windsor, ontariomeasurement and modeling for human exposure assessment. *Environmental Research*, 106(1):7–16, 2008.

[43] K. Whitehouse and D. Culler. Calibration as parameter estimation in sensor networks. In *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pages 59–67. ACM, 2002.

[44] D. Williamson, R. Parker, and J. Kendrick. The box plot: A simple visual method to interpret data. *Annals of Internal Medicine*, 110(11):916–921, 1989.

[45] World Health Organization. Ambient (outdoor) air pollution in cities database 2016. http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/, September 2016. [Online; accessed on 26-September-2016].

[46] Y. Zhang, H. Su, L. Zhong, Y. Cheng, L. Zeng, X. Wang, Y. Xiang, J. Wang, D. Gao, and M. Shao. Regional ozone pollution and observation-based approach for analyzing ozone–precursor relationship during the PRIDE-PRD2004 campaign. *Atmospheric Environment*, 42(25):6203–6218, 2008.