

Low-cost Sensor Array Devices as a Method for Reliable Assessment of Exposure to Traffic-related Air Pollution

by

Nataliya Mykhaylova

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Chemical Engineering and Applied Chemistry
University of Toronto

© Copyright by Nataliya Mykhaylova 2018

Low-cost Sensor Array Devices as a Method for Reliable Assessment
of Exposure to Traffic-related Air Pollution

Nataliya Mykhaylova

Doctor of Philosophy

Department of Chemical Engineering and Applied Chemistry

University of Toronto

2018

Abstract

The exposure to air pollutant mixtures is a well-known risk factor for inducing and increasing the severity of diseases. Combining low-cost sensors in arrays holds great potential for real-time monitoring of pollutant exposure because of versatility and aptitude for tracking multi-pollutant exposure. While many low-cost air pollution monitoring devices have been proposed, several underexplored opportunities remain, including sensor-derived pollution indices, source analysis and exposure assessment.

A thorough investigation of different low-cost commercial gas and particulate matter sensors from 5 manufacturers has been conducted and best-performing sensors were identified. A device for monitoring air quality has been developed and tested. Each device consists of an array of commercially available metal oxide semiconductor for monitoring NO_x and O₃, CO, CO₂ and optical sensors for monitoring PM_{2.5}. Level of pollutant exposure has been characterized at different locations in Toronto over 3 different campaigns between 2013 and 2016. These deployments allowed long-term sensor performance to be evaluated under different meteorological conditions as well as different ranges of pollutant concentrations.

Analysis of a large range of gas sensors revealed several key challenges, including high intra-sensor variability, interference from temperature and nonlinearity. Air quality health index estimation from sensor readings was successfully demonstrated. Three aspects of device reproducibility were evaluated:

drift over time, impact of interferences and impact of site-specific mixtures. Three categories of approaches for improving sensor accuracy and reproducibility were tested: nonlinear calibration models, variable transformations and training data selection. Model reproducibility, ability to adjust for multiple combinations of interferences and ability to resolve sites was improved when devices were calibrated at multiple sites. Analysis showed that both short-term and long-term temporal patterns could be resolved and compared at different sites. Background subtraction helped further emphasize the differences and rank sites in terms traffic-related pollution.

Acknowledgments

The author would like to acknowledge the continued support and mentorship of Greg Evans. He is able to instill incredible passion for the field in everyone around him and without his guidance and effective leadership this work would not be possible. The author is also thankful for the wonderful and supportive research team and committee members. The continued support and professional development opportunities offered by AllerGen NCE were also very much appreciated and were key to the success of this work.

Finally, the author would also like to acknowledge the support of her parents Tetyana and Sergiy Mykhaylov as well as Gregory Szilagyi.

Table of Contents

Acknowledgments	iv
Table of Contents.....	v
List of Tables.....	x
List of Figures	xii
List of Equations.....	xxii
List of Abbreviations	xxiii
1 Introduction	1
1.1 Background and motivation.....	1
1.2 Motivation	5
1.2.1 Increasing temporal and spatial resolution	5
1.2.2 Detection of Multiple Pollutants.....	5
1.2.3 Cost reduction.....	6
1.2.4 Research Hypotheses	6
2 Literature review.....	7
2.1 The challenges and opportunities associated with the low-cost sensors	8
2.1.1 Gas sensors	9
2.1.2 Particulate matter sensors	12
2.1.3 Sensor arrays.....	13
2.1.4 Summary	14
2.2 Development and evaluation of calibration models for low-cost AirSENCE devices	15
2.2.1 Calibration methods	15
2.2.2 Performance metrics	16
2.2.3 Pollutant model performance	17

2.2.4	Air quality index estimation performance	18
2.2.5	Summary	19
2.3	Reliability of sensor device calibration models.....	19
2.3.1	Impact of meteorological factors.....	19
2.3.2	Impact of cross-sensitivity	20
2.3.3	Impact of drift	20
2.3.4	Summary	21
2.4	Approaches for improving calibration model performance	22
2.4.1	Physical and operational approaches	24
2.4.2	Statistical approaches.....	24
2.4.3	Impact of training data	26
2.4.4	Drift compensation	26
2.4.5	Summary	27
2.5	Evaluation of temporal patterns using low cost AirSENCE devices	28
2.5.1	The importance of temporal pattern analysis.....	28
2.5.2	Summary	29
2.6	Evaluation of spatial patterns using low cost AirSENCE devices.....	29
2.6.1	The importance of spatial pattern analysis	29
2.6.2	Summary	30
2.7	Evaluation of pollution sources using low cost AirSENCE devices.....	31
2.7.1	Low cost AirSENCE devices as a method to characterize air pollution sources	31
2.7.2	Summary	33
2.8	Literature review summary.....	33
3	Methodology.....	36
3.1	AirSENCE device design.....	36

3.1.1	Enclosure and air circulation system design	38
3.1.2	Electrical design	40
3.2	Calibration, data acquisition and data processing.....	45
3.2.1	Calibration process	46
3.2.2	Data acquisition and processing	47
3.3	Field study procedures.....	49
3.3.1	Spatial Characteristics of Ultrafine Particles in Toronto (SCULPT) Campaign and long-term performance study.....	52
3.3.2	Pan Am study	52
3.3.3	Multi-site validation study.....	53
4	Results and discussion	55
4.1	Sensor performance.....	55
4.1.1	Intra-sensor variation is a big challenge	55
4.1.2	Variation in sensitivity	58
4.1.3	Influential factors.....	60
4.1.4	Sensor multicollinearity and distribution	62
4.1.5	Relative sensor importance.....	64
4.1.6	Site variation	67
4.1.7	Sensor response drift.....	69
4.1.8	Summary	73
4.2	Pollutant model performance	73
4.2.1	Multivariable models	74
4.2.2	Model performance.....	76
4.2.3	Accuracy and precision.....	80
4.2.4	Model nonlinearity affects performance	82

4.2.5	Inter-device performance variation	83
4.2.6	Multi-pollutant patterns.....	85
4.2.7	Summary	88
4.3	Reproducibility & the impact of interferences	88
4.3.1	Pollutant model drift analysis.....	91
4.3.2	Interference impact	94
4.3.3	Influential factors	96
4.3.4	Site variation	98
4.3.5	Summary	100
4.4	Performance optimization	100
4.4.1	Nonlinear models.....	102
4.4.2	Transformed sensor importance	103
4.4.3	Variable transformation impact.....	104
4.4.4	Training site impact	106
4.4.5	Summary	108
4.5	Temporal patterns.....	109
4.5.1	Temporal pattern estimation accuracy	110
4.5.2	Spatial variation of temporal patterns.....	111
4.5.3	Source contribution analysis	114
4.5.4	Summary	117
4.6	Spatial patterns.....	117
4.6.1	Absolute spatial pollutant variation.....	118
4.6.2	Validation of differences between sites	120
4.6.3	Impact of time on spatial variation	123
4.6.4	Impact of wind on spatial variation	125

4.6.5	Summary	127
4.7	Traffic-related pollution	127
4.7.1	Local vs background impact	129
4.7.2	Major source identification	133
4.7.3	Traffic source comparison	137
4.7.4	Summary	144
5	Conclusions	146
6	Future Work	150
7	References.....	153
8	Appendices	168
8.1	Appendix 1: Commercially-available air monitoring devices	168
8.2	Appendix 2: Design for the AirSENCE sensor device	170
8.2.1	Electrical design for the AirSENCE sensor device.....	170
8.2.2	Code for AirSENCE sensor device.....	178
8.2.3	Mechanical system design for the AirSENCE sensor device	196
8.3	Appendix 3: Finalized calibration model predictor subsets and coefficients.....	199
8.4	Appendix 4: Precision and accuracy analysis	199
8.5	Appendix 5: Wind impact analysis.....	202
8.6	Appendix 6: CO/NO _x ratio analysis.....	203

List of Tables

Table 1. Common air pollutants and their associated sources, concentration ranges and health effects. It is common for pollutants to be emitted by several sources and to be present in complex mixtures making interferences a problem for cross-responsive sensors (adapted from Williams and Kilaru 2014).....	2
Table 2. Common sensing and transduction methods used for air pollution monitoring.....	9
Table 3. Three key sensor aspects governing gas monitoring using metal oxide sensors.....	11
Table 4. Advantages and disadvantages associated with MOS sensors.	12
Table 5. Methods for improving sensitivity and selectivity of gas sensors.....	22
Table 6. Application requirements for air quality monitoring devices (adapted from Williams and Kilaru 2014).	32
Table 7. Key attributes of AirSENCE device design.....	37
Table 8. AirSENCE device revision description.	41
Table 9. An overview of the data sets collected using AirSENCE devices during field studies.	51
Table 10. Sensor subset included in each AirSENCE device.	56
Table 11. Best-performing sensor vs pollutant matrix after evaluating variable importance. Pollutants of interest are listed along top row. Selected sensors are shown in green. Sensors that were found to be important as indicated by all 3 analysis methods for each pollutant were used as final predictor subsets.....	67
Table 12. Average, minimum and maximum pollution levels at the three sites investigated during AirSENCE deployment campaigns (April – December 2015).....	68
Table 13. The impact of baseline and range (sensitivity) drift over time on the response of all the tested sensors.	72
Table 14. Calibration performance metrics for inorganic gases and particulate matter. Precision and accuracy were evaluated using pollutant concentration range excluding the data below LOQ. Average value for each of the performance metrics is shown for the set of 4 devices as well as standard deviation, which represents variation between devices. (* LOQ values were evaluated using two methods described in the text). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.	79

Table 15. Standard deviation values associated two Air Quality Health Index (AQHI) estimation methods at SOCAAR, Highway and Island sites. AQHI sensors v1 refers to AQHI MLR model calculated directly using the sensor data, while AQHI sensors v2 refers to AQHI measurements derived from using AirSENCE estimates for O ₃ , NO ₂ and PM _{2.5} in the AQHI equation. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	87
Table 16. Summary of select commercially-available air monitoring devices.....	168
Table 17. Evaluated sensor list.....	175
Table 18. Bill of materials for components used in design.....	176
Table 19. Calibration model coefficients for each of the target pollutants for a typical AirSENCE device. A mode of five AirSENCE devices is shown for each predictor. Data source: Fall 2015, AirSENCE Rev2.....	199

List of Figures

Figure 1. The new paradigm of air quality monitoring.	7
Figure 2. Schematic of MOS sensors commonly used for monitoring gases. Gases get adsorbed on semiconductor sensor grain boundary surface, catalytically activate by heated substrate, and undergo a redox reaction. Redox reaction leads to electron transfer and change in metal oxide conductivity.....	10
Figure 3. Common reactions between metal oxide semiconductor gas sensors and detected gases.	10
Figure 4. Schematic of optical light scattering sensors commonly used for monitoring particulate matter (PT = photodetector; IRED = IR LED).	12
Figure 5. Trade-off between data quality, deployment density and cost (inspired by Snyder et al. 2013).....	34
Figure 6. Overview of the three core areas in need for further research.	35
Figure 7. Diagram of AirSENCE device design.	37
Figure 8. Sensor array device setup during the SCULPT 2013 summer campaign.....	40
Figure 9. System level diagram of AirSENCE device design.	41
Figure 10. AirSENCE device design iterations.....	43
Figure 11. Sensor array device calibration setup.	47
Figure 12. Data processing pipeline. Black colour denotes steps that are always part of the analysis, blue colour denotes steps that are only relevant to some type of analyses. Orange colour denotes data sources.	48
Figure 13. AirSENCE device testing campaigns.	51
Figure 14. Selected sensor array device locations during SCULPT deployment. Devices from deployment 2 (outlined by yellow circles) were selected for analysis of spatial variation of pollution. Device setup during deployment is also shown in the photo from one of the sites.	52
Figure 15. Selected sensor array device locations during Summer 2015 Pan Am Games Campaign. Photos of device setup during deployment are also shown.	53
Figure 16. Sensor array device locations during Winter 2015 validation at Highway 401, SOCAAR and Halans Point. Photos of device setup during deployment are also shown.	54
Figure 17. Inter-device sensor variation of sensor response. Five AirSENCE devices were compared based on raw voltage readings rescaled 0-1. %RSD across the five devices was derived	

for each 30-minute observation and the average %RSD along with the standard deviation of these average %RSD values are shown. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.58

Figure 18. Inter-device sensor variation for the TGS822 sensor. TGS822 is a MOS sensor for volatile organic compounds, such as solvents. The units of the data shown are arbitrary (sensor voltage measurements rescaled 0-1). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.59

Figure 19. The impact of sensor response correction on reducing inter-device sensor variation. Three AirSENCE devices are compared. The units of the data shown are arbitrary (sensor voltage measurements rescaled 0-1). Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.59

Figure 20. Inter-device variation of tgs2620 sensor and the impact of temperature. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1..... 61

Figure 21. Summary of variables contributing to increasing inter-device %RSD. For example, 15 of the 16 sensors were influenced by temperature. Only consistent trends that apply across all sensor replicates are shown..... 62

Figure 22. Assessment of strength of correlation between sensors. Pearson correlation coefficients are shown only when bigger than 0.5. The quantity of stars indicates significance as indicated by p value. The sensor data is shown in arbitrary units. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2..... 63

Figure 23. Assessment of strength of correlation between sensors and reference monitoring instruments for carbon monoxide, carbon dioxide, nitric oxides, ozone and particulate matter. Pearson correlation coefficients are shown. The quantity of stars indicates significance as indicated by p value. The sensor data is shown in arbitrary units, while reference data is shown in concentration units (i.e. ppb, ugm-3). Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2..... 64

Figure 24. Variable selection process. 67

Figure 25. Pearson correlation coefficients between reference instruments and AirSENCE devices for NO_x and CO at three contrasting sites. The numbers shown are Pearson correlation coefficients multiplied by 100. The colour of ellipses and the shape indicates the degree of correlation. The negative sign indicates negative correlation. Original data was collected

between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.....	69
Figure 26. The process for evaluating sensor drift.....	70
Figure 27. Change in baseline (drift) and range (sensitivity) in P/N706 sensor response (average of four devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	71
Figure 28. Comparison of drift for sensor with high sensor drift (P/N706) and low sensor drift (MQ3). Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.	72
Figure 29. Sensor baseline drift correction for P/N706 using MQ3 sensor data. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.....	72
Figure 30. The performance of ozone MLR model improves as predictors are added to the model. The output of the ozone sensor model $[O_3]''$ is compared to a model making use of both ozone and NO _x sensor predictors $[O_3]'$ as well as a model combining ozone, NO _x , temperature and humidity sensors $[O_3]$. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	75
Figure 31. Evaluation of accuracy and precision of ozone model with change in ozone concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of ozone concentration over the ozone concentration range during the calibration week. Error bars represent variation between devices in predicting ozone concentration (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of ozone concentration. Vertical lines represent limit of detection (LOD). Error bars represent variation between devices in predicting ozone concentration (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.	81
Figure 32. Model nonlinearity and its effect on pollutant models. Time series plot of NO _x MLR model output indicates nonlinearity at higher NO _x levels (A) and % nonlinearity calculation suggests that NO _x and O ₃ are affected by nonlinearity the most (B). Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	83

Figure 33. Comparison for inter-device variation and reference residual. Data for five devices is shown. The regression models were calibrated and tested at the same site using the same reference air pollution monitors. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.	85
Figure 34. Analysis of residuals between sensor device MLR models and reference instrument (A) compared to residuals between sensor device MLR models and an average sensor MLR model for 5 devices (B). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	85
Figure 35. Comparison of performance of two Air Quality Health Index (AQHI) estimation methods at SOCAAR, Highway and Island sites. AQHI sensors v1 refers to AQHI MLR model calculated directly using the sensor data, while AQHI sensors v2 refers to AQHI measurements derived from using AirSENCE estimates for O ₃ , NO ₂ and PM _{2.5} in the AQHI equation. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	87
Figure 36. The process for analyzing the impact of drift and interferences on pollutant model performance.	91
Figure 37. Impact of time on pollution baseline and range measured by AirSENCE devices. Impact is expressed as change over time (difference between end of Spring 2014 and Fall). Both baseline and range declined over time as indicated by negative % change in baseline and range. Average of 3 device replicates is shown. In addition, comparison of inter-device variation (%RSD) for NO _x , O ₃ , CO, PM _{2.5} and CO ₂ is shown as a function of time. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.	93
Figure 38. Drift results in inaccurate long-term assessment of multipollutant patterns. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	94
Figure 39. Impact of inferences on NO _x sensor response, expressed as the difference in strength of the correlation between sensor device output and reference instruments ($R^2_{\text{Pearson NOx sensor}} - R^2_{\text{Pearson NOx reference}}$). Person correlation coefficient was calculated between sensor device NO _x and reference measurements for NO _x , CO ₂ , O ₃ and CO ($R^2_{\text{Pearson NOx sensor}}$). The extent of error in the correlation strength was then quantified by subtracting the Pearson correlation coefficient for reference NO _x and reference measurements for NO _x , CO ₂ , O ₃ and CO ($R^2_{\text{Pearson NOx reference}}$). The value of $R^2_{\text{Pearson NOx reference}}$ for reference NO _x vs. reference measurements for NO _x was 1.0. Average of 6 AirSENCE devices is shown. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.	95

Figure 40. Impact of pollutant concentration on inter-device variation. Comparison of inter-device variation (%RSD) for NO _x , O ₃ , CO, PM _{2.5} and CO ₂ is also shown as a function of pollutant concentration. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	97
Figure 41. Impact of temperature and humidity on pollution baseline and range measured by AirSENCE devices. Impact is expressed as difference between lowest and highest quartiles. Positive change indicates increase with the factor. Average of 3 device replicates is shown. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	98
Figure 42. NO _x model performance before (top) and after (bottom) deployment at the highway site. Average weekly trend over a month is shown. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.....	99
Figure 43. Analysis of improvement in performance when nonlinear calibration models are tested. Comparison of MLR and SVM models for monitoring PM _{2.5} at the Highway 401 site is shown. All models were trained at SOCAAR location. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.....	103
Figure 44. Best-performing sensor transforms.	104
Figure 45. Baseline subtraction for NO _x measurements at site 18. Baseline is indicated in orange. One-minute resolution data is shown.	105
Figure 46. Comparison of non-transformed and transformed O ₃ regression model performance tested at the island and highway sites. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.....	106
Figure 47. Comparison of CO calibration model trained at SOCAAR and tested at Highway 401 and Island sites with CO regression model trained at the sites and tested at SOCAAR. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	108
Figure 48. Evaluation of the ability of AirSENCE devices to resolve short-term and long-term variation in pollutant levels. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	111

Figure 49. Comparison of NO _x temporal variation for AirSENCE devices at sites 12, 18, 10, 15, 5 and central site (1 min average data). Average of 4 weekdays is shown. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	113
Figure 50. Evaluation of the ability of AirSENCE devices to compare temporal patterns of O ₃ at 3 different sites. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	114
Figure 51. Comparison in differences in diurnal pattern for NO _x at site 18. Traffic level was measured in volume of vehicles per day within a 1500m buffer. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	115
Figure 52. Evaluation of the ability of AirSENCE devices to accurately access temporal variability of CO relative to NO _x . Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	116
Figure 53. Comparison of the absolute NO _x level differences measured by AirSENCE devices (shown in box-and-whisker distribution plots) and those derived using land use regression models calibrated from 2002 fall and 2004 spring monitoring campaigns with 143 reference monitors (shown in a map). Sensor data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	120
Figure 54. Assessment of the performance of CO measurement at SOCAAR, Highway 401 and Island sites. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	121
Figure 55. Comparison of distributions of NO _x at the island, SOCCAR and highway sites for regression models derived using transformed and raw data. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.	123
Figure 56. Analysis of the impact of time of day on divergence between central site and other sites. Summary of the similarity between central site and other sites in the morning is also shown. Kolmogorov-Smirnov D statistic is used as a metric of divergence. Morning is defined as 5 AM-11 AM. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	124

Figure 57. Analysis of the variation in divergence between sites depending on time of day and day of the week. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.....	125
Figure 58. Analysis of the variation in divergence between sites for NO _x and CO depending on wind direction. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.....	126
Figure 59. Pollutant levels estimated by AirSENCE devices vary over the course of the day and depend on the amount of traffic. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.	130
Figure 60. Pollution component contributions (inspired by Fuzzi et al. 2015).....	131
Figure 61. Analysis of local differences between site 15 (site with higher traffic influence) and site 5 (site with lower traffic influence) sensor measurements (1 min average data). Original sensor measurements are compared to those with background removed. *Traffic data is shown in vehicles per day within a 1500m buffer. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.....	132
Figure 62. Comparison between local pollution source contribution and the traffic volume at different sites. *Traffic data is shown in vehicles per day within a 1500m buffer. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	133
Figure 63. Comparison of pollutant signatures associated with different sites. Axes represent absolute Pearson correlation coefficients between NO _x and other pollutants estimated by AirSENCE devices at different sites (central site, site 18, and site 5). A larger shaded area indicates stronger correlation. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	134
Figure 64. Variation in pollution source contribution over a course of a day. The graph is split into panels to show the dominant wind direction and speed for different times of day: morning, afternoon, evening and night. The NO _x _1m colour scale indicates the NO _x concentration measurements by the sensor device. Wind speeds are split into the intervals shown by the scale in each panel. The grey circles show the % frequency occurrence based on 1 min data. The plots show the proportion (here represented as a percentage) of time that the wind is from a certain angle and wind speed range. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.....	136

Figure 65. Variation in traffic volume at SOCAAR in the morning, afternoon and evening.....	136
Figure 66. Variation in pollution source contribution over a course of 5 months. The graph is split into panels to show the dominant wind direction and speed for different months. The NO _x _lm colour scale indicates the NO _x concentration measurements by the sensor device. Wind speeds are split into the intervals shown by the scale in each panel. The grey circles show the % frequency occurrence based on 1 min data. The plots show the proportion (here represented as a percentage) of time that the wind is from a certain angle and wind speed range. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.....	137
Figure 67. Differences in traffic emissions associated with gasoline and diesel vehicles.....	138
Figure 68. Assessment in variation of local source contribution for 3 traffic-related pollutants: NO _x , CO and CO ₂ . Ranking of sites based on local traffic influence and comparison of 3 traffic-related pollutants is also shown. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.....	139
Figure 69. Comparison of diurnal trends for CO ₂ , CO and NO _x at site 18 for original data (A) and background-corrected data (B) (1 min average data. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	140
Figure 70. Time series of local contribution to NO _x and CO levels at site 18 (1 min average data). Comparison of relative levels of CO and NO _x for weekday and weekend are shown. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.	142
Figure 71. Diurnal and weekly trend clusters for 4 types of sensors (inorganic gas, carbon monoxide, organic gas and particulate matter) at 4 different sites are shown. Raw sensor voltage output was normalized at each site prior to comparison. Only relative changes displayed are meaningful, not absolute levels. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.....	144
Figure 72. Summary of the key contributions of the thesis to the field of research.	149
Figure 73. AirSENCE devices installed on TTC streetcars for monitoring ozone, carbon monoxide, carbon dioxide, and PM _{2.5} along King Street routes.	151
Figure 74. System Level Diagram of Electrical Components.	171
Figure 75. The design for circuit boards comprising data logging unit as well as images of completed and assembled boards.....	172

Figure 76. Data logging unit circuit schematics.....	172
Figure 77. The design for the power interface board as well as image of a completed and assembled board.....	173
Figure 78. Power interface board circuit schematics.	173
Figure 79. The design for one of the sensor boards as well as image of a completed and assembled board.....	174
Figure 80. Sensor board circuit schematics.....	174
Figure 81. A sketch of mechanical design as well as implemented device and image of deployment at a site.	198
Figure 82. Two small versions of the device enclosure were designed and built. Version 1 was designed to be included in backpack used to monitor the exposure of runners. Version 2 contains a bigger range of sensors and was designed to fit into the lung transplant patient exposure assessment kit.	198
Figure 83. Evaluation of accuracy and precision of NO _x model with change in NO _x concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	200
Figure 84. Evaluation of accuracy and precision of CO model with change in CO concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	200
Figure 85. Evaluation of accuracy and precision of CO ₂ model with change in CO ₂ concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.....	201

Figure 86. Evaluation of accuracy and precision of PM_{2.5} model with change in PM_{2.5} concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.201

Figure 87. Assessment of the impact of wind direction on NO_x concentration measured by AirSENCE devices and reference instruments. The graph is split into panels to show the dominant wind direction and speed for different ranges of reference NO_x observations (Nox_Out): very low, low, medium and high. The NO_x_lm colour scale indicates the NO_x concentration measurements by the sensor device. Wind speeds are split into the intervals shown by the scale in each panel. The grey circles show the % frequency occurrence based on 1 min data. The plots show the proportion (here represented as a percentage) of time that the wind is from a certain angle and wind speed range. Mean values in each panel indicate the mean sensor NO_x output for each panel (each bin of reference NO_x observations). Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.....202

Figure 88. The comparison of CO/NO_x ratios (medians and distributions) for different sites (1 min average data). Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.204

List of Equations

Equation 1. Equation for precision or inter-device variation. %RSD is % relative standard deviation and s is the standard deviation of the sensor device replicate measurements. For precision measurement, \bar{x} is the average pollutant estimate for a concentration level for one device, while for inter-device variation, \bar{x} is average pollutant estimate for sensor device replicates.	77
Equation 2. Accuracy equation. A is average absolute % Error, A_t is the target pollutant reference value, F_t is sensor measurement and n is the number of observations.	78
Equation 3. Sensitivity equation. M is sensitivity, Δy is change sensor device regression model output, Δx is change in pollutant concentration during the same period.	78
Equation 4. Nonlinearity equation. %NL is percent best fit nonlinearity, D_{max} is maximum output deviation from the best-fit line and FS is full-scale output.	79
Equation 5. Air Quality Health Index calculation equation.	86

List of Abbreviations

ANN = Artificial Neural Network

AQHI = Air Quality Health Index

CI = Confidence Interval

CO = Carbon Monoxide

CO₂ = Carbon Dioxide

MCU = Microcontroller unit

MLR = Multiple Linear Regression

MOE = Ontario Ministry of the Environment

MOS = Metal oxide semiconductor

NO = Nitric Oxide

NO₂ = Nitrogen Dioxide

NOX = Oxides of Nitrogen

O₃ = Ozone

Pan Am = Pan American

PCB = Printed Circuit Board

PM = Particulate Matter

PM_{2.5} = Particulate Matter with diameter ≤ 2.5 microns

PM₁₀ = Particulate Matter with diameter ≤ 10 microns

RH = Relative Humidity

RMSE = Root Mean Square Error

rs = Spearman's Rank Order Correlation Coefficient

SCULPT = Spatial Characteristics of Ultrafine Particles in Toronto

SE = Standard Error

SOCAAR = Southern Ontario Centre for Atmospheric Aerosol Research

SVM = Support Vector Machine

T = Temperature

VOCs = Volatile Organic Compounds

WDir = Wind Direction

WHO = World Health Organization

WS = Wind Speed

1 Introduction

1.1 Background and motivation

The quality of the air we breathe is a cause for concern for environmental and public health researchers alike (Fuzzi et al. 2015). More than 3.2 million deaths per year have been attributed to air pollution (Jerrett 2015). This value can be even greater for young children, the elderly, and people with disabilities (Gordon et al. 1999; UNICEF 2016; Buka et al. 2006). Awareness of chemical pollution and its influence on human health and the environment has given rise to increased regulatory control on emissions of toxic compounds into the environment, forcing both industry and academia to respond with effective monitoring and control measures. The interest in a more complete understanding of the health effects of environmental pollutants and pollutant mixtures has also led to a rising demand for robust studies on their distribution with increasing spatial and temporal resolution. The ability to compare pollutant levels and types in different microenvironments is also becoming increasingly important. These trends have prompted vigorous research and the development of methods and devices for sensitive monitoring of multiple airborne pollutants.

Furthermore, it is difficult to disregard the multi-chemical and multi-source nature of human exposure to airborne pollutants (Table 1). These effects are particularly important when the mixture components influence each other's response. Hamada et al have suggested that fine particulate matter may facilitate exposure to toxic vapours (Hamada et al. 2000). NO₂ has been shown to increase the inflammatory effect of aeroallergen exposure in asthma patients (Barck et al. 2005). A report commissioned in the beginning of 2010 by the Environment Directorate General of the European Commission has demonstrated that health implications are greater for pollutant mixtures than for the sum of their components, even at trace concentrations (Kortenkamp, Backhaus, and Faust 2009). Sensitive multi-pollutant analysis can be accomplished using sampling-laboratory analysis protocols. Most of the studies contributing to the body of knowledge in this field have used a sampling-laboratory analysis protocol (Corrêa et al. 2010; Araki et al. 2010; Cindoruk and Tasdemir 2010; Ramadan 2010; Zhang et al. 2010; Chang et al. 2010). These analyses are intrinsically off-line, time-consuming, expensive, and have been linked to problems with sample integrity and artifacts.

A review of the literature in the field of air quality monitoring and exposure assessment research yields several recurring areas of interest:

1. Detection of a diverse range of pollutants;
2. Improvement in sensitivity of detection of trace levels of pollutants;
3. Quantitation and correction for chemical interactions and interferences;
4. Improvement of specifications (i.e. miniaturization, cost reduction and ease of operation).

Table 1. Common air pollutants and their associated sources, concentration ranges and health effects. It is common for pollutants to be emitted by several sources and to be present in complex mixtures making interferences a problem for cross-responsive sensors (adapted from Williams and Kilaru 2014).

Air pollutant	Type	Source Example	Typical Range	Regulatory threshold level ¹	Health Effects ²
Ozone (O ₃)	Secondary and sometimes primary	Photo-chemical via NO _x and organics; photocopy, laser machines	0-150 ppb	62 ppb (8 hours)	-Breathing problems -Reduced lung function -Asthma -Eye irritation
Carbon monoxide (CO)	Primary	Fuel combustion – mobile sources,	0-0.3 ppm	9 ppm (8 hours) 35 ppm	-Respiratory diseases

¹ CCME developed Canadian Ambient Air Quality Standards (CAAQS; <http://airquality-qualitedelair.ccme.ca/en/>) for PM_{2.5}, O₃, and NO₂. For particulate matter, ambient standards are shown as 3-year averages of the annual 98th percentile of the daily 24-hour average concentrations or annual average of all 1-hour concentrations. For ozone, the 3-year average of the annual 4th highest daily 8-hour average concentrations is used, while for nitrogen dioxide, the 3-year average of the annual 98th percentile of the daily maximum 1-hour average concentrations and average over a single calendar year of all 1-hour average concentrations are used. EPA National Ambient Air Quality Standards (40 CFR part 50; <https://www.epa.gov/criteria-air-pollutants/naaqs-table>) are shown for CO since corresponding ambient standards are not available in Canada. No regulatory exposure limits are available for CO₂.

² Adapted from U.S. Environmental Protection Agency. Air Pollution Monitoring. Available online: <http://www.epa.gov/oar/oaqps/montring.html>. Additional data from Air Quality Ontario, Ministry of the Environment, <http://www.airqualityontario.com/>

		industrial processes		(1 hour)	-Cardiovascular diseases
Nitrogen dioxide (NO ₂)	Primary and Secondary	Fuel combustion – mobile sources, electric utilities, off-road equipment	0-50 ppb	42 ppb (1 hour) 12 ppb (1 year)	-Respiratory diseases -Cardiovascular diseases
Carbon dioxide (CO ₂)	Primary	Fuel combustion – electric utilities, mobile sources	350-600 ppm	None	-Fatigue
Fine particulate matter (PM _{2.5})	Primary and Secondary	Fuel combustion (mobile sources, electric utilities, industrial processes), dust, agriculture, fires	0-40 µg/m ³	27 µg/m ³ (24 hours) 8.8 µg/m ³ (1 year)	-Eye, nose, throat irritation -bronchitis -lung cancer

Many innovative technologies that have been recently proposed for air pollution monitoring and exposure assessment (Choi et al. 2010; DeVito et al. 2009; Khatko et al. 2009; Penza et al. 2008; Talazac et al. 2001; Jerrett et al. 2017). Their compact size, fast response, low-cost and wide range of detectable pollutants enable concurrent monitoring at multiple locations, portable monitoring applications as well as more informative monitoring of wide range of pollutants in ambient air. These properties are highly beneficial for environmental exposure studies. However, while the low-cost commercial sensor technologies appear very promising in laboratory-based tests, their performance in complex environments encountered in field studies remains unsatisfactory (Lewis and Edwards 2017). Interference effects from temperature, humidity and chemical reactivity are some of the major challenges to acquiring accurate data for air quality analyses. In addition, the selective detection of a diverse range of existing and emerging compounds of concern in a single measurement has yet to be effectively accomplished. The inherent trade-off between selectivity and the reversibility of sensor surfaces further complicates the design of an optimum air quality monitoring device. However, there are several opportunities for overcoming these challenges. In particular, combining low-cost sensors into arrays would allow a diverse range of air contaminants to be monitored in a single measurement and in real time, while keeping overall monitoring costs low. Monitoring of multiple air contaminants creates

opportunities to use signal processing to lower the impact of variability in physical and chemical interferences on the signal. The use of sensor array architecture can also improve long-term data reliability since parallel data output from multiple sensor replicates can significantly increase statistical power of the measurement.

It is hypothesized that an inexpensive device can be developed, which would utilize time-resolved data for airborne pollutant mixtures to provide an integrated measure of exposure. The first objective of this PhD project was to develop a modular portable system for the analysis of pollutant mixtures at ambient levels. An array of commercial sensors was combined into a real-time data collection system with built-in data storage. The second objective of this PhD project was to assess the performance of this system in field studies. The portable, inexpensive multi-pollutant monitoring system developed during this project was expected to facilitate the comparison of the temporal profiles of major pollutants in different microenvironments, allowing their exposure profiles to be developed.

This thesis document is organized as follows:

- Chapter 1 summarizes the background and motivation for undertaking the proposed research. The importance of increasing temporal and spatial resolution of monitoring, parallel detection of multiple pollutants and cost reduction is outlined.
- Chapter 2 provides a synthesis of relevant literature, including gaps and opportunities for extending the capabilities of low cost sensor technologies as well as potential applications.
- Chapter 3 provides an overview of methodology of the research including the design of AirSENCE devices, data analysis methodology, calibration process and studies that were carried out.
- Chapter 4 outlines 7 key areas of research, including analysis of sensor performance, pollutant model performance, the impact of interferences, performance optimization methods, analysis of temporal and spatial patterns as well as analysis of sources, such as traffic.
- Chapter 5 summarizes the findings and conclusions.
- Chapter 6 provides an overview of the future work.

1.2 Motivation

The primary goal of this research was to show that low-cost devices for simultaneous, multi-pollutant measurement can be designed to provide measurements of urban air pollution with acceptable accuracy and precision. It was hypothesised that the advantages of sensor array approach combined with the multivariate signal processing techniques would mitigate challenges associated with low-cost sensor performance.

1.2.1 *Increasing temporal and spatial resolution*

One of the greatest priorities in pollution exposure research is the ability to acquire high time and spatial resolution data. Urban air pollution exposure is characterized by highly variable time-activity patterns and air pollution is known to be highly variable both temporally and spatially (Steinle, Reis, and Sabel 2013). Therefore, high time resolution monitoring is essential for exposure assessment (Kumar et al. 2016). Therefore, a key goal of this project was to design a low cost (below \$500) and portable device capable of obtaining high time resolution data. Device portability and low cost increase the number of devices that can be deployed, thereby increasing the spatial resolution of the measurements and allow for a more accurate exposure assessment and source analysis. Fast sensor response, combined with the multi-pollutant data, would allow for a more comprehensive assessment of the temporal and spatial patterns in pollutant concentrations.

1.2.2 *Detection of Multiple Pollutants*

The second key goal of this research project was to improve the ability to measure and resolve multiple pollutants within air pollution mixtures. The need to analyze as many air pollutants as possible is balanced by the availability, the cost, and the long-term robustness of sensors. The following major urban pollutants are the focus of this research: NO_x, PM_{2.5}, CO, CO₂, O₃. Although the 1-hour time resolution data is available at 4 Ministry of Environment and Climate Change air quality monitoring stations in Toronto for NO_x, O₃ and PM_{2.5}, their spatial coverage of spatially heterogeneous pollutants is limited. NO_x and CO concentrations can be highly variable in space and a knowledge of this variability is essential to our understanding of the impacts of combustion-related pollution sources. Measurement of O₃ is key for understanding pollutant chemistry. CO₂ is important for understanding of emission factors.

PM_{2.5} measurements improve our understanding of pollution background and source contribution. Temperature and relative humidity are important for understanding the impact of meteorology on pollutant levels. Furthermore, most of both gases and particulate matter can have serious impact on population health and these impacts often vary in space and time (Jerrett 2015). Concurrent monitoring of multiple ambient air contaminants has many potential benefits such as the improved data quality, more thorough exposure assessment, improved characterization of pollution sources and pollution gradients throughout a city.

1.2.3 Cost reduction

The development of novel low-cost integrated circuit technologies and manufacturing processes have contributed to a range of low cost commercially available sensors. In addition, open source hardware and software toolkits have enabled reduction in product development costs. Furthermore, wireless technology enabled continuous monitoring and remote device adjustment that further reduced air pollution study costs. The requirement for trained personnel is one of the highest long-term costs associated with air quality monitoring studies. Hence, another major goal was the design of a system that is robust and can be operated reliably for extended periods without human intervention. To this end, the priorities included: 1) quantification of the accuracy of the sensor response, 2) development of a method to correct the sensor response for short-term noise and long-term drift, and 3) wireless networking for data acquisition and remote diagnostics. The low cost of the commercially available sensors and microcontrollers as well as wireless signal acquisition and remote device diagnostics were all key drivers of cost savings.

1.2.4 Research Hypotheses

The overarching research hypothesis for this thesis was as follows: “Coupling multiple sensors (sensor array devices) with appropriate data mining methods can yield pollutant concentrations that reveal meaningful spatial and temporal patterns arising from to traffic emissions”. This hypothesis can be broken down to three supporting hypotheses:

1. Sensor array devices show acceptable performance relative to gold standard instruments.
2. Sensor array devices are able to reveal meaningful temporal and spatial patterns.
3. Sensor array devices can reveal relative contribution of sources such as traffic.

2 Literature review

Holistic understanding of air pollution as well as its variation in space and time is essential for development of effective strategies to reduce the impact of air pollution on our health and environment (Tham 2016). Traditionally, air quality monitoring is primarily guided by the need to validate regulatory compliance as well as understand and model air pollution sources, sinks and transformation processes. The data is collected by standard monitoring equipment that has been well-characterized by both the manufacturer and researchers for a wide variety of applications. More, recently, the paradigm has shifted, driven by community efforts in using a wide range of low-cost, heterogeneous, little-characterized sensors across the world (Aberer et al. 2010; Carton 2015; Snyder et al. 2013). Availability and low cost of a wide range of sensors has prompted many researchers to develop their own equipment. For example, web cameras have been re-proposed for monitoring air quality (Wong et al. 2007). Research literature suggests that by combining low-cost AirSENCE devices with standard data and new algorithms, we would be able to derive new insights (Figure 1). The efforts of improving intrinsic sensor performance should be combined with development of generalizable calibration strategies and data processing methods through collaborative work of experts in sensing technology, software development, data processing and pollution monitoring domain expertise (Theunis, Stevens, and Botteldooren 2017).

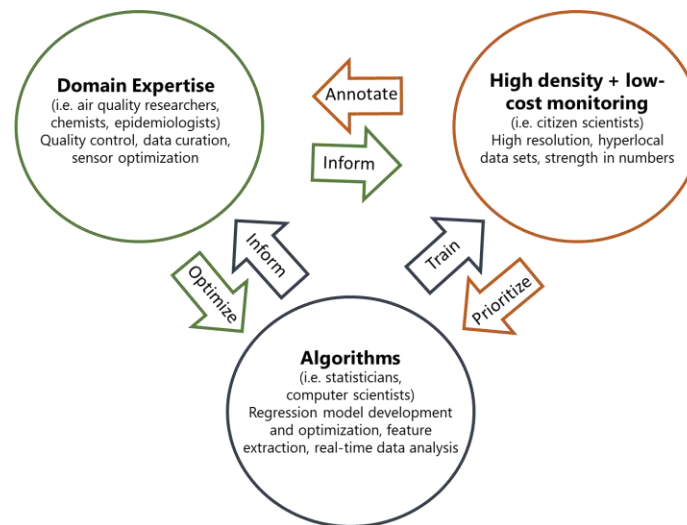


Figure 1. The new paradigm of air quality monitoring.

The field of air quality monitoring using low-cost sensors can be sub-divided into 7 major research topics. The first core area of research is the design and characterization of the sensor modules. The aim is to characterize the performance of the sensors from material, chemical or physical perspective to gain deeper understanding of the underlying processes governing the sensor response and performance. The second core research area is the development and evaluation of regression and classification calibration models relating the sensor response to pollutant types and levels. The third major area of study is the validation of the pollution models by testing their reliability over extended periods of time and under a range of environmental conditions and pollutant levels. The shortcomings identified by the validation studies often lead to proposals of approaches for improving calibration model performance through adaptive models, which can be characterized as another research area. The sensors are then applied to understand 2 major aspects of air pollution exposure research areas: 1) spatial distribution of air pollution, and 2) variation of pollution in time. Finally, spatial and temporal variation in sensor response are combined to understand pollution sources. The focus of this thesis is to address the research gaps in the 7 research topics described above.

2.1 The challenges and opportunities associated with the low-cost sensors

Various sensor technologies have become available over the last 10 years to measure the air quality, but many accuracy and reliability challenges still exist (Wetchakun et al. 2011; Fine et al. 2010; Kanan et al. 2009; Yamazoe 2005). Most of the research developments have been focused on analysis of the criteria pollutants, including carbon monoxide, ozone, nitrogen dioxide and particulate matter (Commercially-available air monitoring devices). Sensor technologies commonly used for detecting these pollutants include metal oxide semiconductors (MOS), electrochemical cells (EC), non-dispersive infrared sensors (NDIR) and light scattering sensors (Table 2). Other technologies such as ultraviolet absorption sensors, light absorption sensors and direct mass transduction sensors have also been used but are not readily available commercially. Each of the sensor technologies comes with tradeoffs with respect to performance and usability. However, metal oxide semiconductor (MOS) sensors and optical light scattering sensors are the least expensive technologies for detecting gas and particulate matter contaminants. Since reducing device cost was a key design goal, MOS and optical sensor technologies were selected for device development and optimization of their performance was a core focus of the thesis.

Table 2. Common sensing and transduction methods used for air pollution monitoring.

Detection method	Common target pollutants	Advantages	Disadvantages
Metal oxide semiconductor (MOS)	CO, O ₃ , NO _x , VOCs	Small size, stable, inexpensive	Sensitive to change of RH, T; cross-sensitivity; high power consumption.
Electrochemical cell (EC)	CO, O ₃ , NO _x	Low power, more sensitive and linear than MOS	Large impact from interferences and drift, smaller lifetime
Non-dispersive infrared (NDIR)	CO ₂	Compact, stable	Sensitivity depends on path length; calibration is easy to misinterpret
Light scattering	PM	Small, inexpensive, easily available commercially	Not a direct mass measurement; cannot detect ultrafine particulate matter

2.1.1 Gas sensors

A common low-cost method for monitoring is metal oxide semiconductor (MOS) sensors (Figure 2). MOS sensors respond to gases through the change in resistance of a semiconductor film in response to oxidation and reduction reactions that occur on its surface (Figure 3). The MOS sensor surface becomes catalytically active in response to heating, which facilitates adsorption, reaction and desorption of gases.

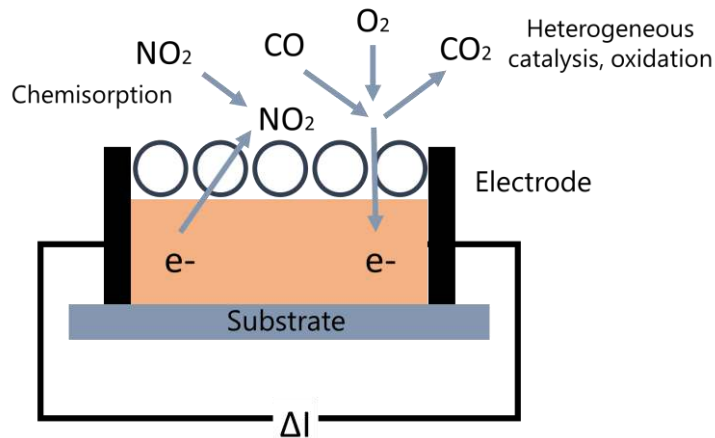


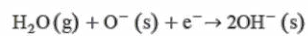
Figure 2. Schematic of MOS sensors commonly used for monitoring gases. Gases get adsorbed on semiconductor sensor grain boundary surface, catalytically activate by heated substrate, and undergo a redox reaction. Redox reaction leads to electron transfer and change in metal oxide conductivity.

Clean air reactions

Oxygen

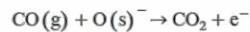


Water vapour

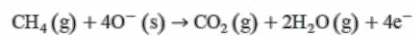


Reducing gas reactions

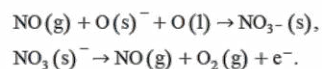
Carbon monoxide



Volatile organics

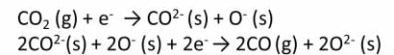


Nitric oxide

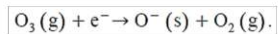


Oxidizing gas reactions

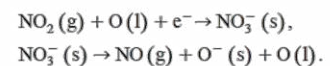
Carbon dioxide



Ozone



Nitrogen dioxide



Nitric oxide

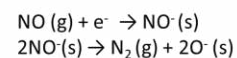


Figure 3. Common reactions between metal oxide semiconductor gas sensors and detected gases.

The change in MOS sensor resistance is the result of three types of processes: 1) sensing process (controlled by gas-solid interaction), 2) electrical transduction process (controlled by microstructure), 3) gas diffusion process (controlled by the sensor structure and manufacture properties) (Table 3).

Table 3. Three key sensor aspects governing gas monitoring using metal oxide sensors.

Sensor surface properties	Transducer properties	Sensor structure properties
Gas adsorption	Carrier mobility	Pore size
Redox reactions	Grain size	Diffusion depth
	Doping	Film thickness

Some of the ongoing challenges associated with MOS sensors are sensor variability and poor performance of the devices in field. While sensors were shown to have acceptable performance for some pollutants (i.e. ozone), others have been found to have poor correlation to reference monitors (Theunis, Stevens, and Botteldooren 2017). Variation between sensors of the same type is generally low, but the sensor cross-sensitivity is a major problem that often cannot be explained by known correlation between pollutants because it depends on a variety of environmental factors.

Another challenge in using the metal oxide semiconductor sensors is that the sensor response and sensitivity is impacted by the order of gas exposure (Sarry and Lumbreras 1999). This is particularly significant when the gases have opposite effect on sensor response (i.e. oxidizing vs reducing).

Novel semiconductor sensors have recently been proposed that have better response to oxidizing gases and less dependence on humidity and better stability, but are still in research stage (Korotcenkov 2007). While the choice of metal oxide material generally affects both sensing and transduction processes, new composite sensors that allow for optimization of both processes separately have been proposed (Korotcenkov 2007). Studies suggest that detection of ambient pollutant mixtures using metal oxide sensors is still challenging due to noisy response and cross-sensitivity (Williams, Kaufman, and Garvey 2015). Drift over time is another challenge for metal oxide semiconductor sensors. Some researchers proposed methods to adjust drifted sensor response using other sensors which have correlated response but which show less drift (Aguar et al. 2015). The sensors were used as a supplementary tool to standard monitors to assess urban traffic but could not be used independently because of large cross sensitivity. The authors suggest that multi-sensor array combined with field calibration would help correct for interferences. A summary of the advantages and disadvantages associated with MOS sensors, according to the literature, is provided in Table 4. A major focus of the thesis was to validate the advantages and overcome the disadvantages through improved device design and data analysis techniques.

Table 4. Advantages and disadvantages associated with MOS sensors.

Advantages	Disadvantages
<ul style="list-style-type: none"> • More responsive to low levels of pollutants • Fast response time once stabilized • Good repeatability • Long sensor life • Least expensive technology • Responds to large range of compounds • Small size • Can operate at higher temperature & pressure 	<ul style="list-style-type: none"> • Sensitive to interferences • Not linear at higher concentrations • Impacted by drift • Slow start-up due to warm-up • Shorter battery life due to heated sensor element

2.1.2 Particulate matter sensors

The most common sensor used for detection of particulate matter is optical scattering sensor, which works by measuring the light (commonly IR) refracted by the particulate matter passing through an optical chamber (Figure 4).

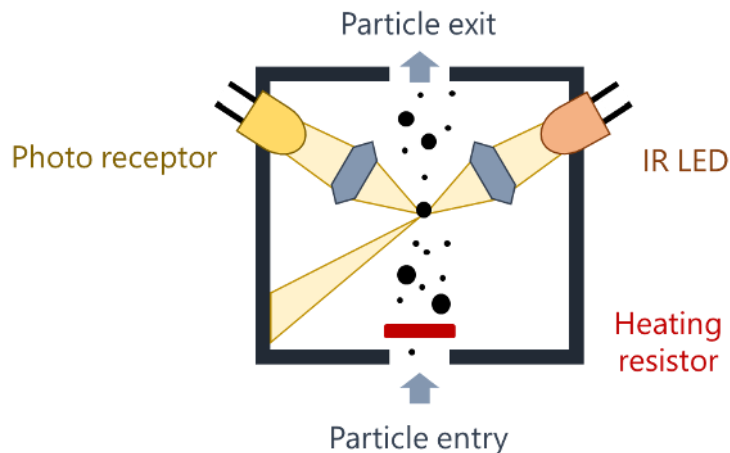


Figure 4. Schematic of optical light scattering sensors commonly used for monitoring particulate matter (PT = photodetector; IRED = IR LED).

Very few commercial particulate matter monitors are capable of accurately assessing the variation in particulate matter concentration (Williams et al. 2014). While some monitors have acceptable performance (i.e. Dylos DC1100, Met One, RTI MicroPEM), they are too expensive and too large for many applications. A few sensors have been shown to be able to detect pollution hot spots, but the sensitivity was too low for long-term air quality monitoring (Budde, Busse, and Beigl 2012). Sensors are still limited by the high cost and power requirements (Jovaseevic-Stojanovic et al. 2015). Studies show that field calibration could yield better results than interpolation of reference instrument data, although extensive device-specific and site-specific calibration was required (Holstius et al. 2014). More work is needed to better characterize variation between sensors and operation parameters. While some laser-based optical scattering sensors are correlated with research-grade equipment, the accuracy varies with particle properties and the type of sampling used (Kelly et al. 2017). The particle size distribution and particle composition have large effect on sensor signal, resulting in low correlation with particle mass concentration. More work is needed to develop accurate models for complex aerosol systems. In addition, a large impact of humidity exists due to hygroscopic growth of particles that leads to overestimation of particle mass concentration. More inter-comparison studies with standard instruments in specific particle mixtures are required due to variation in sensor design, mode of particle transfer and the impact of humidity (Manikonda et al. 2016). Other research studies showed that the performance depends on particle composition, size, ambient humidity, resulting in difference in sensor output by as much as a factor of 10 (Wang et al. 2015). Modifying the airflow system, particle concentration algorithm and averaging interval helped improve the performance. Sensors were demonstrated to have linear response on average, with some saturation at higher particulate matter concentrations and higher relative standard deviation with decrease in concentration. Field studies also suggest that low-cost optical aerosol sensors are impacted by the temperature interference and the drift in sensor baseline response by more than 30% in three weeks (Olivares and Edwards 2015). More work is needed to understand sensor response towards different aerosol compositions and size distributions as well as understand the sensor variability and the causes for drift and temperature interference.

2.1.3 Sensor arrays

One of the key opportunities highlighted by US Environmental Protection Agency in recent report on low cost sensor technology is the use of sensor arrays to measure multiple pollutants in parallel, the need for more field studies, quality control and techniques for improving sensor sensitivity, selectivity and

reliability over extended periods of time (MacDonnell et al. 2014). Combining an array of cross-responsive sensor with multivariate statistics has been suggested as a particularly promising technique for assessing air pollution variation and sources (Schutze 2008).

The optimum performance of sensor arrays depends on the selection of best combinations of sensors and sensor feature predictors. The selection of sensors depends on multiple parameters, including sensitivity, selectivity, hysteresis, stability over the time and on exposure to other gases in the environment. In addition, the choice of sensor operation temperature can determine chemical, structural, and long-term stability. While lower temperatures are associated with higher sensitivity and longer lifetime, higher temperatures lead to lower dependence on humidity, better signal reproducibility as well as fast response and recovery (Korotcenkov 2007). Feature selection methodologies have been used to reduce the level of drift and noise, leading to more accurate, lower-cost and power-efficient devices (Gardner, Boilot, and Hines 2005; Gualdrón et al. 2007). Studies have shown that subsampling features can reveal optimum combination of features for specific applications (Nowotny et al. 2013). It was found that while some sensor technologies outperform others when subsets are compared, the combination of the different sensor technologies in a single array can lead to the best performance. Studies have also suggested that more accurate and robust models can be built through better feature selection methods. A cascaded genetic algorithm was used to select variables to identify and quantify volatile organic compounds detected by metal oxide semiconductor sensor array (Llobet et al. 2004). The method also had an added advantage of reduced computation time due to a smaller subset of the best-performing variables selected. Feature selection was used to identify new features that show superior performance compared to ratio between sensor resistance in gas and clean air (R_s/R_0) suggested by most sensor datasheets and was found to improve the classification and lead to a better understanding of the data (Pardo and Sberveglieri 2007).

2.1.4 *Summary*

Low-cost sensors and sensor arrays have been shown to have promising characteristics for applications in air quality studies. However, several research gaps remain. Sensor datasheets contain very limited, and often misleading information on sensor performance as well as information that was collected exclusively under lab conditions. In addition, no published studies, to the best of the author's knowledge, provide comprehensive comparison of sensor performance across a range of commercial metal oxide semiconductor sensors in ambient air environment. The impacts of environmental

conditions and pollutant concentration on sensor performance and the variability between sensors of the same type need to be characterized in more detail. The variability between sensors has a large impact on calibration model performance, but its characteristics and causes have not been explored in the literature in a comprehensive manner. In addition, the extent of correlation of sensor response between different types of sensors and between sensors and pollutants has not been described in the literature. The characterization of correlations (i.e. spread, slope, intercept) can give insights into the cross-responsiveness of sensors, how it varies with environmental conditions and the suitability of specific sensors as predictors for target pollutants. In addition, literature review has not revealed any studies characterizing sensor response distribution. A detailed analysis of sensor distribution is needed to better understand the dynamic range of sensor response and how it varies with environmental conditions.

2.2 Development and evaluation of calibration models for low-cost AirSENCE devices

2.2.1 *Calibration methods*

As indicated earlier, inherent variability of low-cost sensors makes it necessary to conduct extensive calibrations prior to using the devices for monitoring air quality. Several recent studies have compared different calibration approaches and proposed new and innovative calibration methods that save costs and improve performance (Appendix 1:). The performance of collocation calibration and lab calibration was compared for a custom low-cost air quality monitor for CO, O₃, NO₂, total VOCs, and CO₂ (Piedrahita et al. 2014). In addition to being time-consuming and expensive, lab calibration was found to exhibit higher measurement error due to narrow environmental variable space that can be generated under lab conditions, differences in composition of lab and ambient air as well as meteorological variability in ambient environments. However, collocation calibration also had some limitations, including lack of control over gas interferences, pollutant range and variation in temperature and humidity. These factors resulted in flatter and noisier calibration curves. The calibration of individual sensors and testing under a range of environmental conditions and air pollution mixtures in field was found to improve the performance. Model performance and assessment of diurnal trends was further improved when the calibration location and time were optimized to ensure the higher variability in the training dataset. Similar results have also been shown with volatile organic sensor array, where calibration with wider range of gas concentrations can lead to more a generalizable performance in response to new gas

mixtures (Rodriguez-Lujan et al. 2014). A semi-supervised calibration technique has been shown to help improve the performance of sensor array at classifying field samples of carbon monoxide, formaldehyde, and toluene samples, while reducing the time and cost of training the sensor array (Huang et al. 2016).

Adaptive calibration approaches, where the calibration model parameters are continuously adjusted, have also been tested by several research groups. A continuous re-calibration approach for enhancing the quality of pollution measurements has been proposed and validated in field using ozone sensor measurements in a public transportation vehicle-mounted sensor network (Hasenfratz, Saukh, and Thiele 2012). The approach was found to significantly reduce sensor aging, cross-sensitivity as well as reduce the measurement error by a factor of ten. It was also least dependent on the amount of reference data, which is important for mobile networks where the location of reference stations and the mobile node schedule cannot be changed. On-the-fly calibration correction method exploiting periodic proximity between devices as well as reference monitors, was found to stabilize device readings and improve accuracy (Budde et al. 2013). However, more work is needed to automate sensor calibration and reduce the influence of environmental factors such as humidity and air pressure. A set of approaches related to adaptive calibration are calibration transfer approaches, where the calibration model is transferred between devices. Calibration transfer between devices can reduce cost, save time, improve mass production efficiency as well as reduce the requirement for recalibration of drifting sensor. However, large amount of mismatch between the response of sensor replicates makes the calibration transfer difficult. Several different calibration transfer methods were evaluated for gas sensor arrays (Fonollosa et al. 2016) and Direct Standardization was demonstrated to have the best performance both for transferring calibration between two different devices to reduce calibration costs and for the same device over time to mitigate drift.

2.2.2 Performance metrics

A major challenge in ensuring high performance air pollution monitoring at low cost is development of a generalizable set of metrics that can be used to validate device performance. Mean error and correlation to standard equipment are commonly used metrics to assess sensor performance, but they do not predict sensor performance in large-scale field applications and do not cover all aspects of possible differences between sensors and standard equipment as well as changes in performance over time. Other groups proposed an entirely new set of criteria that allows for a more holistic performance

assessment for a larger set of applications and environments (Robinson et al. 2017). The suggested criteria include 8 different metrics: 1) Root Mean Squared Error (RMSE), 2) Pearson correlation coefficient, 3) Kandel correlation coefficient, 4) Spearman correlation coefficient, 5) ability to locate pollution sources, 6) ability to represent the pollution level on a coarse scale, 7) ability to capture the high temporal variability and 8) metric reliability in space and time. The criteria were deployed for assessment of diverse set of gas and particle sensors across 8 cities in Europe and provided a unique information on sensor performance as well as a system to identify systematic and sensor-specific faults. For sensor arrays, the performance metrics can be more difficult to quantify since the component sensors often have different performance characteristics but are considered in combination for pollutant concentration assessment. A new approach for defining the limit of detection (LOD) for multi-sensor systems has been found to provide more accurate comparisons between sensor systems by using a broader definition of noise (Fonollosa et al. 2014). The method considers the variability of the noise distribution with pollutant concentration and generates LOD for different concentration levels. The method is more generalizable to different types of noise and better captures the information contained in the analytical signals and prior knowledge, allowing for more accurate benchmarks of analytical systems.

2.2.3 *Pollutant model performance*

Several different low-cost air pollution monitoring devices have recently been described in the literature. Appendix 1 highlights some of the commonly-used commercially-available devices along with their performance metrics. Most devices based on electrochemical sensors for NO_2 and O_3 have been shown to have good linearity, detection limit under 10 ppb and response time under 2 minutes (Spinelle, Gerboles, and Aleixandre 2015). However, high cross-sensitivity between NO_2 and O_3 as well as humidity and long-term drift have significant effect on sensor response. A drift of over 6 ppb over a period of 3 months was observed. Combining the multiple sensor measurements with multivariate regression method had a positive effect on sensor performance, allowing for the data quality objectives of the European Directive for indicative methods to be achieved (measurement uncertainty of 30%). While some device sensors (i.e. CairClip NO_2) have good linearity, repeatability and lack hysteresis, the sensor was shown to suffer from long-term drift and ozone interference (Spinelle, Gerboles, and Aleixandre 2013). Long-term ambient air calibration was needed to develop reliable models to estimate NO_2 . On the other hand, Cairclip O_3 / NO_2 sensor did not suffer from long-term drift. However,

calibration in field over a period of one week was still necessary for reliable performance. Field calibration was also found to be essential for measuring volatile organic mixtures using sensor arrays (Romain and Nicolas 2007). Recent multi-year, multi-city CAIRSENSE study provided an overview of performance of emerging ambient air quality devices measuring oxides of nitrogen, ozone, carbon monoxide, and particles under real-world conditions (Jiao et al. 2016). The study indicated high variability in correlation to reference monitors ($r = 0.4 - 0.60$) even for devices composed of identical sensors.

2.2.4 Air quality index estimation performance

Air quality indices can be used represent cumulative air quality sensor measurements and exposure. Sensor arrays have a potential to become useful platforms for quantifying air quality indices and their variation in space and time since they can be designed multiple air pollutants at once while maintaining low cost and small size that enable large scale deployment. Several studies have recently explored the performance of sensor arrays for quantifying air quality indices. Low-cost multi-sensor devices were integrated in distributed city-network and used to assess spatial variation in pollution exposure and EPA air quality index (AQI) (Penza et al. 2014). Long-term accuracy of the index as well as the individual NO_2 , CO and PM_{10} pollutant measurements was tested. The study demonstrated good performance of the air quality index estimates but a significant cross-sensitivity between nitrogen dioxide and ozone. The NO_2 sensor correlation coefficient was 0.79 for the correlation with NO_2 analyzer and 0.83 for the combination output of O_3 and NO_2 analyzers.

EPA air quality index (AQI) often tends to underestimate the air pollution and is subjective since it does not consider all health hazards (Sowlat et al. 2011). A novel air quality index based on fuzzy logic has been proposed to overcome these limitations and combine the criteria pollutant data (i.e. CO, SO_2 , PM_{10} , O_3 , NO_2) with organic pollutant data (benzene, toluene, ethylbenzene, xylene, 1,3-butadiene) as well as differential weighting factors. According to the case study conducted over a period of one year at 5 different sampling stations, the index based on fuzzy inference proved to be a useful tool for ranking of air quality. A new Multipollutant-Multisite Air Quality Index has also been proposed as a better framework allows the comparison of air quality over time and across sites (Plaia et al. 2013). The index is derived through a two-step aggregation, using spatial PCA synthesis and pollutant synthesis across CO, NO_2 , PM_{10} and SO_2 pollutants. The index was demonstrated to have multiple advantages, including

ability to account for the coexistence of multiple air pollutants, more consistent performance across space and time and reduced ambiguity (false positive).

2.2.5 *Summary*

In summary, a large range of air quality monitoring devices as well as regression methods and calibration approaches have been proposed in the literature and by commercial companies within the last 10 years. However, only limited data is available on the performance of models as a function of sensor type and number. In addition, more work is needed to compare the calibration models for different pollutants with respect to their accuracy, precision, linearity, inter-device variation and other calibration performance metrics. Analysis of causes for performance limitations would help identify opportunities for improvement. The comparison of performance between device replicates and with changes in pollutant concentration would also reveal insights regarding the best calibration approach for the AirSENCE devices. No research studies have analyzed the accuracy of multi-pollutant health indices (e.g. Air Quality Health Index), and their reliability with changes in environmental conditions. Most of the air quality studies that exist rely on calculating the index from the estimated pollutant concentration. Deriving the pollution indices using multivariate regression could lead to lower error and improved performance compared to calculation approach. However, to the best of author's knowledge, no studies exist comparing the two approaches.

2.3 Reliability of sensor device calibration models

2.3.1 *Impact of meteorological factors*

Impact of interferences can be a big challenge for accurate pollution monitoring. Some of the major factors influencing sensor response are meteorological factors, including changes in temperature and humidity (Williams and Kilaru 2014). For example, ozone sensors calibrated in a laboratory conditions were found to suffer from changes in temperature, the influence of which declined when the devices were calibrated with ambient air during field study (Gerboles and Buzica 2009). Humidity was also demonstrated to have large reducing effect on response of most commercial metal oxide sensors (Delpha, Siadat, and Lumbreras 1999). The assessment of CO, NO₂ and O₃ model performance has

indicated high sensor variability due to weather conditions and emission patterns (Castell et al. 2017). A linear dynamic model was proposed that describes the change metal oxide semiconductor sensor resistance as a function of gas concentration under known changes in temperature and humidity (Hirobayashi, Kimura, and Oyabu 1999). The effect of humidity is similar for different types of tin oxide sensors and can be approximated by exponential function for steady state response and linear function for dynamic response. Beyond just changes in ambient environment, the physical design of sensor device can exacerbate the impact of temperature and humidity on device performance. The air flow, material temperature dissipation and even sensor orientation can create issues with overheating, overcooling on condensation, which can impact sensor performance. Methods have been proposed for detecting anomalies in temperature distribution and correcting for them by optimizing the PCB and 3D-printed case design, thereby improving device accuracy, by reducing the error to 2% (Salamone et al. 2017).

2.3.2 Impact of cross-sensitivity

Cross-sensitivity is another major factor influencing sensor performance. Good performance in the laboratory is not indicative of a good performance under real-world conditions (Castell et al. 2017). Due to low measurement repeatability, individual sensor calibration as well as rigorous evaluation under a range of environmental conditions are needed. In many cases, interferences lead to a greater sensor response than the target pollutant (Lewis et al. 2016). While trace gases (i.e. CO₂) generally have small effect on sensor response, the high concentration of these gases relative to target pollutant can still lead to large artefacts. Supervised machine learning approaches were shown to have better performance than standard individual parameter regression in correcting for the effect of interferences (Lewis et al. 2016). Challenges remain in development effective strategies for sensor data quality control. A comprehensive calibration and testing in the presence of co-pollutants and interferences is important to correct for sub-optimal measurement but is lacking in the literature.

2.3.3 Impact of drift

While sensor arrays are often used to counteract selectivity and calibration issues, they require a period of calibration co-located with reference device to develop a calibration model. One of the key challenges is that the resulting model is often specific to calibration site and season, while gas composition varies

between sites and seasons, resulting in deteriorated performance over time. Temperature variation has been shown to be one of the factors contributing to sensor drift, by introducing shifts in dynamic features of sensor resistance and decreasing sensor stability (Abidin, Asmat, and Hamidon 2013). The changes in ambient and operating temperature were used to detect initial stages of sensor drift and its compensation. The impact of different causes on sensor degradation was compared for 8 different types of commercial metal oxide sensors incorporated in sensor arrays (Fernandez, Marco, and Gutierrez-Galvez 2015). Electrical failure resulted in least degradation in sensor performance, followed by sensor poisoning, while sensor ageing had the worst effect on performance. Electrical failure and poisoning had the effect of bringing the sensor response to a fixed level, while sensor poisoning also introduced random noise. Both sensor redundancy and diversity were important for a sensor device to maintain discrimination capacity and be less prone to damage.

2.3.4 *Summary*

While temperature and humidity have been shown to have a large impact on sensor performance, the difference in their impact on performance parameters, such as sensitivity, baseline and inter-device variation has not been explored in detail. Most of the existing studies that explore the impact of temperature and humidity on sensor performance do not consider their impact in long term and in the context of ambient air. It is important to characterize the types and magnitude of interferences as well as their impact on observation uncertainty. Quantification of the change of specific performance metrics for each environmental factor and their relative impact on accuracy can help better understand the various components of sensor response, calculate measurement confidence to flag the faulty measurements as well as develop methods for correcting the sensor response. Comparison of the pollutant models with respect to the influence of interferences in ambient air can also help understand the expected level of uncertainty for each pollutant for a specific type of site as well as the return on investment on adding a capability of monitoring an air contaminant to a device. The magnitude of the impact of interferences can also help predict how well a device will perform at a site with different pollutant mixture. The accuracy of calibration models should be validated at several sites with different levels of air pollution and types of air pollutant mixtures. Validation of device accuracy at several sites as well as ranking of models based on the extent to which their performance is generalizable to different combinations of interferences can help better prioritize the sensor maintenance and analysis. Another key aspect of sensor device model performance that needs further exploration is drift. While drift in baseline has been explored in several research studies, the drift in sensitivity has received

relatively little attention in the literature. In particular, more studies are needed for quantifying the rate of drift over time to better understand when the devices need to be re-calibrated or replaced. The evaluation of the impact of drift on regression model parameters would help develop correction factors to extend the useful lifetime of sensors in field. In addition, the evaluation of the impact of drift on the inter-device variation can help understand the extent to which spatial variation can be quantified by sensor device networks.

2.4 Approaches for improving calibration model performance

Three types of approaches have been proposed in the literature for improving sensitivity and selectivity of AirSENCE devices: 1) physical modification approaches, 2) operational approaches, 3) statistical approaches (Table 5) (Theunis, Stevens, and Botteldooren 2017; Kumar et al. 2016). Physical modification approaches include gas pre-concentration, removal of interfering gases through scrubbers and filters and development of sensor arrays. Operational approaches include thermal cycling and modulation of the airflow speed. Statistical approaches include sensor response transformation, the use of nonlinear regression models, reconstruction of sensor signal from noise and optimization of the training dataset. Several approaches have been proposed in the literature for improving the sensitivity and selectivity of sensors (Barsan, Koziej, and Weimar 2007; Liu et al. 2012) but this area is still a subject of ongoing research.

Table 5. Methods for improving sensitivity and selectivity of gas sensors.

Approach	Impact on sensitivity	Impact on selectivity
Physical modification approaches		
Pre-concentrator (an add-on to sensor chamber that provides controlled adsorption and release of gases)	Enables an increase in target gas concentration	Pre-concentrator can be designed to be selective
Scrubber (an add-on to sensor chamber that removes	Removal of interferences can improve signal-to-noise ratio for detection of target analyte	Removal of interferences improves detection selectivity

select gases from the air)		
Sensor array (a combination of cross-responsive sensors)	Sensor redundancy can improve sensitivity by increasing the signal-to-noise ratio	Provides multi-dimensional signatures; corrects for variation in temperature and humidity; improves overall performance when target gases have different optimum sensing conditions
Operations approaches		
Thermal cycling (operation of sensors using a controlled heating cycle)	Improves sensitivity for gases with different optimum sensing temperatures	Improves selectivity for gases with different optimum sensing temperatures
Modulation of airflow speed (control of airflow through the sensor chamber)	Faster air exchange range improves sensitivity towards fast-changing sources	-
Statistical approaches		
Signal transforms	Improve linearity and sensitivity for a greater range of pollution concentrations	-
Nonlinear regression models	Improve linearity and sensitivity for a greater range of pollution concentrations	Can improve performance through interference correction
Reconstruction of data from noise	Improves sensitivity by reducing the impact of interferences (baseline shifts, cross-sensitivity, aging) and improves performance	Improves selectivity by reducing the impact of interferences (baseline shifts, cross-sensitivity, aging) and improves performance
Training dataset optimization	Improves sensitivity by training on a dataset with a greater range of pollutant concentration levels	Improves selectivity by training on a dataset with a greater range of interferences

2.4.1 *Physical and operational approaches*

The sensitivity and selectivity of metal oxide gas sensors can change significantly when sensors are operated at a different temperature (Fonollosa et al. 2013). The modulation of working temperature of metal oxide sensors has been a popular method for improving the accuracy of discriminating and quantifying gas mixtures using sensor arrays, while reducing the effects of noise and nonlinearity (Vergara et al. 2005). Periodic heating and cooling of the metal oxide surface alters the kinetics of diffusion and reaction processes of the gas on sensor surface. Combining temperature modulation with global feature extraction method has led to improved device accuracy, reduced measurement time and power consumption (Vergara, Martinelli, et al. 2007; Vergara, Llobet, et al. 2007). Novel sensing approach that makes use of a single gas sensor and an algorithm that classifies sensor response kinetics has been proposed. The system can identify chemicals accurately regardless of their concentrations and is able to estimate concentrations of the analytes, while reducing the complexity of hardware and software. A new methodology was proposed for selecting the optimal operation temperature for each sensor in a gas sensor array based on a combined index of performance (Fonollosa et al. 2013). The method provided precise information on the ability of the sensor array to discriminate organic compounds based on the temperature of component sensors. New metal oxide sensor active sampling microchambers were proposed as a method to better control gas depletion reactions to identify individual pollution species in gas mixtures and detect ambient pollutant levels (Becker et al. 2000). In addition, the design of sensor chamber and air sampling system were found to impact the device performance. Smaller sensor chambers with hydrophobic walls, higher operation temperatures and active sampling were shown to increase the speed of analysis and sensitivity while allowing for calibration and self-test features that reduce drift and improve performance. The design of better calibration and validation protocols that acknowledge both the advantages and the limitations of commercial sensors has been highlighted by several recent publications as the subject of future work (Lewis and Edwards 2017).

2.4.2 *Statistical approaches*

Several interesting approaches have been proposed in the last few years for sensor optimization using statistical methods. In particular, nonlinear regression models have recently become more common for processing sensor data. The performance of linear, multilinear and artificial neural network (ANN)

calibration models was compared for a sensor array of ozone, nitrogen dioxide, carbon dioxide and carbon monoxide sensors (Spinelle, Gerboles, and Aleixandre 2015). Two-week calibration and five-month testing revealed that simple multi-linear regression was sufficient for ozone estimation, while an ANN model was needed for estimation of NO₂, NO, CO and CO₂. ANN with multiple sensor predictors of mixed types (metal oxide, electrochemical, optical) showed the highest correlation with reference measurement and the lowest bias, while correcting for cross-sensitivity issues. ANN was able to also correct for temperature and humidity interference for NO₂ and CO₂ without the need for temperature and humidity sensors, although these were required for NO and CO models. Neural network ensemble and dynamic ensemble integration have recently been proposed for online concentration estimation of formaldehyde, benzene, toluene, and carbon monoxide using a metal oxide sensor array (Kadri et al. 2013). Sensor fusion algorithms based on neural networks and support vector machines have also been proposed for continuous monitoring and identification of NO₂–NH₃–humid air mixtures (DeVito et al. 2007). Analysis of short term sensor response dynamics was used to correct for poor stability and selectivity of sensors. In a different study, independent component analysis (ICA) was used to differentiate and remove the temperature and humidity influence on the data, leading to 20% improvement in classification accuracy (Di Natale, Martinelli, and D’Amico 2002). The proposed method was also found to reduce disturbance from environmental parameters without directly measuring them. The performance of a two-dimensional classifier ensemble has been demonstrated over a period of 3 years (Liu, Chu, and Tang 2015). The method has been shown to have a good long-term performance and independence of gas concentration. A metal oxide sensor array was combined with a fuzzy-logic algorithm to identify, discriminate and quantify low levels of carbon monoxide and nitrogen dioxide in mixtures with relative humidity and volatile organic compounds, while ensuring selectivity required for indoor air quality applications and HVAC control (Zampolli et al. 2004). A new method for monitoring and forecasting ozone, nitrogen dioxide, and sulfur dioxide using low-cost array of gas and meteorological sensors has been demonstrated (Shaban, Kadri, and Rezk 2016). Three machine learning algorithms were evaluated for accuracy in using historical data to predict pollutant levels 1-24 hours into the future. It was found that the decision tree-based M5P model produced the most accurate and generalizable performance, SVM was adaptable to high dimensional data but showed mediocre performance, while ANN achieved the worst performance because of its tendency to overfit and poor ability to generalize when working with a relatively small training dataset. Multivariate modeling, where pollutant, meteorological and temporal variables were combined, was found to enhance prediction accuracy because of interdependence of the captured variables.

2.4.3 *Impact of training data*

Several studies have also explored the impact of training data and calibration model transfer on regression model performance in field. Neural network regression applied on a metal oxide sensor array was used to detect CO, NO₂, NO_x and benzene (DeVito et al. 2009). Two-week field calibration was sufficient to ensure consistent performance. A different neural network model trained for one week on a sensor array with field data was found to accurately estimate benzene for over 6 months (DeVito et al. 2008). Training period greater than 25 days further improved model performance, reducing benzene overestimation. However, seasonal variation impacted long-term model performance, reducing the model accuracy at low benzene levels. A method for transferring calibration models between metal oxide sensor array devices designed for monitoring volatile organic compounds has been demonstrated (Fernandez et al. 2016). The comparison of four different calibration techniques lead to selection of the Piecewise Direct Standardization method that is least sensitive to the variations in the working temperature and leads to the best predictive performance.

As described in Table 3, sensor signal is often composed of a number of different components. Some researchers have indicated good results with dynamic multivariate modeling approaches. For instance, dynamic neural networks have been combined with multi-sensor devices to reduce the influence of cross-sensitivity on NO₂, NO_x and O₃ concentrations estimated by AirSENCE devices (Esposito et al. 2016). Dynamic neural networks work by exploiting sensor response dynamics and rapid transients to improve accuracy and reduce uncertainty of model output. The method has been shown to achieve significant reduction in estimation error for all target pollutants, particularly when the target gas concentration is rapidly changing. A novel technique has been proposed to reconstruct signal from Poisson noise to compensate for dynamic and systematic errors (Budde, Köpke, and Beigl 2015). The technique is robust against offset, drift and some types of cross-sensitivity. It can also be applied in post-processing and doesn't assume any prior knowledge about time since calibration and the extent of drift. It is therefore suitable for use by non-expert users and analysis of publicly-available datasets. However, some of the challenges of the method include large amount of cross-sensitivity that is similar in frequency to signal and high frequency variation in systematic error.

2.4.4 *Drift compensation*

Drift over time is another major drawback of low-cost sensors and has been an active subject of several different research studies with a number of different drift compensation approaches being proposed in the literature. A simple and robust algorithm has been successfully demonstrated to compensate the drift in an array of commercial metal oxide sensors. The algorithm makes use of curve fitting of the temporal variation of the sensor signal to describe the temporal drift and is suitable for real-time drift compensation applications (Haugen, Tomic, and Kvaal 2000). An energy-efficient data mining algorithm based on random forest ensemble learning and feature selection has been proposed to tracking gas sensor array drift (Alwadi and Chetty 2015). The advantages of the method include lack of assumptions and physical analysis needed as well as adaption to system parameters reducing the need for interrupting system operation during sensor response adjustment. A novel methodology based on multiple kernels was proposed for increasing the robustness of sensor arrays against sensor failures as well as extending sensor array device lifetime (Fonollosa, Vergara, and Huerta 2013). A dynamic drift tracking and reduction system using adaptive estimation algorithms has been proposed and demonstrated to perform well in situations where traditional pattern recognition techniques have failed (Holmberg et al. 1997).

2.4.5 *Summary*

Optimization of AirSENCE devices to improve sensitivity, selectivity and long-term performance is an active area of research. However, there are still several research gaps. While nonparametric regression models have been shown to yield better performance with non-linear sensor response, their training is often computationally expensive and requires a highly variable training dataset. A more detailed analysis is needed to compare the performance of standard parametric regression approaches with alternative regression methods, such as neural networks and support vector machines. In particular, a comparison of parametric and nonparametric approaches for the full range of pollutants of interest and for a range of pollutant levels across different sites would be very beneficial for selection of an optimum data analysis pipeline. In addition, a number of data processing methods are available that improve the performance of nonlinear sensor models while maintaining the analysis simplicity of linear regression. For example, variable transformation approaches are promising, but have received relatively little attention in the literature. In addition, development of multi-component regression, where a separate model is built for different pollutant concentration ranges, is another promising approach that needs to be tested. Furthermore, the following research topics need further exploration: 1) the impact of

training data subset characteristics on calibration model performance, 2) the techniques for adjusting sensor response for interferences, and 3) the techniques for resolving sites during the analysis of spatial variation of pollution. For example, a comparison of the performance of calibration models trained at sites with high, medium and low pollution levels would help design better device calibration protocols. Finally, the extent to which long-term historical training data can be used to improve model performance should be investigated in more detail as this would also be another important consideration for optimization of calibration protocols.

2.5 Evaluation of temporal patterns using low cost AirSENCE devices

2.5.1 *The importance of temporal pattern analysis*

Air pollution is known to be highly variable both temporally and spatially, while urban exposure is characterized by highly variable time-activity patterns (Steinle, Reis, and Sabel 2013). Therefore, high time resolution monitoring is important for reliable exposure assessment (Kumar et al. 2016). Fast response of many low-cost sensor technologies has enabled a variety of new air pollution analysis applications, such as traffic emission factor analysis, commute exposure monitoring and real-time sensor networks (Fuentes et al. 2016). They were also used to enable gas source localization and distribution mapping (Monroy, Gonzalez-Jirenez, and Blanco 2012). High frequency sensor-based monitoring device mounted on an unmanned aerial vehicle was used for identification and 3D mapping of pollution sources (Malaver et al. 2015). One of the limitations of such studies has been the stabilization time of metal oxide semiconductor sensors³. However, studies on transient feature analysis has provided a way to expedite pollutant quantification and classification while ensuring the stable performance (Muezzinoglu et al. 2009). The use of transient features has an added advantage of providing power savings and resulting in predictor features that are less affected by drift and other interferences. The temporal variation patterns detected by sensors have been shown to be key for spatiotemporal deep learning-based air quality prediction models, which have been shown to have superior performance and temporal stability (Li et al. 2016). The time series sensor data combined with temporal pattern visualization techniques were found to reveal unique patterns and relationships

³ After the sensors have been turned off, they need time to stabilize as the metal oxide semiconductor heats up

between air quality, wind direction and other meteorological factors (Du et al. 2017). The changes in temporal patterns have also been used for early failure detection (Wu and Song 2013) and long-term drift prediction and compensation (Zhang et al. 2013). High resolution AirSENCE devices can improve accuracy by identifying outliers, while providing an early warning system to help reduce exposure to air pollution (Yu and Lin 2015). Portable sensor platforms have been proposed for collecting real-time data on public transport and as part of citizen-driven pollution sensing (Devarakonda et al. 2013). While the device temporal resolution was sufficient to track exposure, more work is needed to develop reliable calibration protocols.

2.5.2 *Summary*

As shown above, a number of studies have explored the applications of high time resolution signal processing provided by sensor arrays for air quality monitoring and exposure assessment. However, more detailed analysis of the ability of AirSENCE devices to monitor both short-term and long-term temporal patterns is needed. Temporal patterns are key for understanding the characteristics of sources, their variation in space and their influence on exposure. The variation in peak magnitude, timing and frequency could be informative of local sources and emission factors but is often lost during temporal averaging in air quality studies. Comparing temporal patterns for different pollutants could reveal the characteristics of traffic at the site as well as help validate the accuracy of the device. For example, the analysis of CO relative to NO_x could suggest temporal variation in gasoline, diesel and other types of combustion source emissions. However, such analysis is lacking in the literature.

2.6 Evaluation of spatial patterns using low cost AirSENCE devices

2.6.1 *The importance of spatial pattern analysis*

Research studies have repeatedly shown strong dependence of ambient urban air pollution levels on local emissions and meteorology resulting in strong spatial and temporal gradients (Borge et al. 2016). Therefore, high density air quality monitoring is essential for full understanding of sources and spatial distribution on air pollution. Reviewers suggested that while low-cost sensors enable high spatial resolution monitoring, the cost of sensor network installation, maintenance and data analysis will likely

exceed the cost of AirSENCE devices themselves. Therefore, we also need to develop better approaches for optimizing sensor networks and the design of data analysis pipelines (Kumar et al. 2015).

A number of research studies have explored the use of low-cost sensors in high density networks. A network of low-cost particulate matter sensors was used to identify the city air quality patterns and hotspots, effectively supplementing existing city air quality networks (Gao, Cao, and Seto 2015). A high density continuous air quality monitoring network was used to map air pollution exposure and develop a new algorithm to suggest routes that help reduce exposure (Hasenfratz et al. 2015). The high spatial resolution data was used to develop land-use regression models and continuously improve their accuracy using historical measurements. The intra-site variability and contributions of large emission sources as well as local activities were effectively detected by dense low-cost sensor networks, which were shown to have comparable performance in resolving spatial variation as standard monitor networks (Miskell, Salmond, and Williams 2017). Other studies have demonstrated that while sensor networks can detect spatiotemporal variation at sub-neighborhood spatial scale, frequent field calibrations are required to ensure sensor consistency (Moltchanov et al. 2015). A new method for optimizing sensor networks to improve the precision of air quality monitoring while reducing the cost has also been proposed (Boubrima, Bechkit, and Rivano 2016). The model incorporates sensor drift and weather conditions to select sensor locations such that the quality of air pollution estimation interpolated between the sensor nodes is maintained. The model was evaluated in Lyon city and was found to perform 3 times better than random and uniform deployment.

2.6.2 *Summary*

Assessment of spatial variation of air pollution is one of the most common applications of low cost sensors. However, there are several research topics that would benefit from more attention. First, the spatial variation in pollutant levels measured by AirSENCE devices needs to be validated in a more comprehensive manner during sensor network studies. This can be done by deploying the AirSENCE devices at existing government pollution monitoring sites. In addition, most studies in the literature assess the spatial variation in average levels primarily. However, pollution peak levels (i.e. top 5% of the pollutant concentration distribution) could be more informative of the local sources and provide a better means of comparison between sites. Another key gap in the literature is the lack of availability of comparisons between different types of pollutant models in terms of their ability to resolve differences between sites. The dependence of the differences between sites on the amount of traffic is another key

area of interest, since it is expected that the differences between sites are likely to be mostly caused by local pollution sources. In addition, analysis of the dependence of the differences between sites on the wind direction, time of the day and day of the week could be informative for design of better study protocols. Low-cost sensors are known to suffer from cross-sensitivity and lack of linearity, which could be affecting the analysis of spatial distribution of air pollution. Therefore, another key area that needs further analysis is the impact of nonlinear models and sensor response transformation on the accuracy of spatial variation analysis.

2.7 Evaluation of pollution sources using low cost AirSENCE devices

2.7.1 *Low cost AirSENCE devices as a method to characterize air pollution sources*

Due to the ubiquity of most air pollutants, identifying specific sources and comparing the relative quantities of pollutants emitted from these sources can be challenging. Source apportionment is essential for development of strategies to control and eliminate emissions (Leoni et al. 2018; Zhang et al. 2018). Complex receptor modeling and expensive long-term air quality monitoring studies have been undertaken to attempt linking air pollutants to their corresponding sources (Ou et al. 2018). Sensor arrays can facilitate automated and potentially much simpler strategies for source apportionment through analysis of the pollutant mixture response of the sensor array.

A number of studies explored the use of low cost sensors to classify pollution sources. Vehicle-mounted sensors have been used to detect and identify urban pollution hotspots such as gas pipeline leaks (von Fischer et al. 2017). A mobile sensor device combined with inverse modelling algorithm and urban building models was used to locate outdoor pollution sources with minimal model training (Xue and Zhai 2017). A semiconductor array composed of 9 commercially available sensors was used to estimate pollution severity and classify sites using categories “urban”, “traffic” and “photochemical pollution” with a success rate of 80% (Barakeh et al. 2017). A calibration and normalization method, employing Neural Network and “fuzzy logic” algorithms, was used to overcome drift and inter-array reproducibility issues. Another study has been conducted to evaluate the applications of gas sensors for monitoring indoor air quality events and interventions, but was shown to have limited capacity to resolve complex organic mixtures at low concentration levels (Caron et al. 2016). Studies assessing higher concentration mixtures, such as combustion sources, have indicated more promising results. Analysis showed that differentiation of gasoline or diesel fuel emissions is possible using sensors combined with multivariate

statistics (Schutze 2008). Furthermore, low-cost sensors were used to calibrate city traffic-related air pollution models (Seto et al. 2014). Methods for estimating emission factors using low cost particulate matter sensors across a wide range of pollutant concentrations have been proposed (Johnson et al. 2016). However, additional work is still needed to characterize the sensor performance under a wide range of environmental conditions as well as improve the accuracy in under higher levels of relative humidity.

One of the biggest challenges in air quality monitoring and pollution source analysis studies using low cost sensors is selecting the right sensors and ensuring sufficient data quality for the target application. While regulatory monitoring requires high level of precision and accuracy, useful information can still be generated for supplemental monitoring or hotspot identification application even when the regulatory performance metric goals cannot be achieved (Table 6).

Table 6. Application requirements for air quality monitoring devices (adapted from Williams and Kilaru 2014).

Application area	Pollutants	Precision and Bias Error	Data Completeness ⁴	Rationale (Tier I-IV)
Education and Information	All	<50%	≥ 50%	Collect evidence that the pollutant exists in some wide range of concentration.
Hotspot Identification and Characterization	All	<30%	≥ 75%	Show that the pollutant of interest exists at a concentration that is close to its true value.
Supplemental Monitoring	Criteria pollutants, Air Toxics	<20%	≥ 80%	Provide additional air quality data to complement existing monitors.
Personal Exposure	All	<30%	≥ 80%	Understand factors influencing personal exposures at precision and bias errors that are

⁴ Data Completeness - amount of data that was obtained, compared to the amount that was expected

				representative of those reported in the scientific literature.
Regulatory Monitoring	O ₃ , CO, SO ₂ , NO, PM _{2.5} , PM ₁₀	<7% <10% <15% <10%	≥ 75%	Ensure high quality data is being obtained to meet regulatory requirements

2.7.2 Summary

Several different research groups have proposed new methods for assessment of pollution sources using low cost sensors. Progress has been made in attributing sensor response to various source characteristics. However, there are still several research areas that need to be explored in more detail. The temporal variation in air pollution consists of several different components that can be attributed to different types of sources. However, the feasibility of using the local traffic component as a proxy of local traffic emissions has not been validated. In addition, more research is needed to assess the differences in relative contribution of local component of sensor response relative to background component for traffic related pollutants across a range of sites. This type of information would be useful for ranking sites based of the local pollution and source contribution. Multi-pollutant monitoring capability of sensor arrays is promising for understanding the emission profiles of sources. In particular, correlations between pollutants could reveal source composition, but the use of multi-pollutant correlations as a method to resolve source profile is yet to be explored. Wind direction can also be combined with sensor array data to provide further information regarding the likely local sources. However, few long-term studies made use of wind data to understand the impact of local sources on sensor response as well as compare sites with respect to contribution of traffic emissions. Furthermore, temporal variation in raw sensor response has a potential to reveal the information on pollution sources directly, bypassing the additional errors associated with regression models. However, the feasibility of this approach has yet to be evaluated.

2.8 Literature review summary

There is a lot of potential for high density networks of low-cost sensor devices to drive policy change and improve public health (Reis et al. 2015). Low-cost sensor technologies create capacity for monitoring of more air contaminants at higher temporal and spatial resolution compared to traditional approaches. As a result, they can provide unique insights into the pollution sources and variation of air pollution in

space and time. However, there is often a tradeoff between data quality, deployment density and cost (Figure 5). Some of the major challenges identified include questionable data quality, lack of characterization of commercially available sensors in ambient conditions (Snyder et al. 2013) and lack of detailed sensor drift assessment (McKercher, Salmond, and Vanos 2017). The implementation of quality control procedures would be a key priority moving forward, as is the development of standard guidelines and protocols for sensor evaluations based on their intended applications (Wang and Brauer 2014).

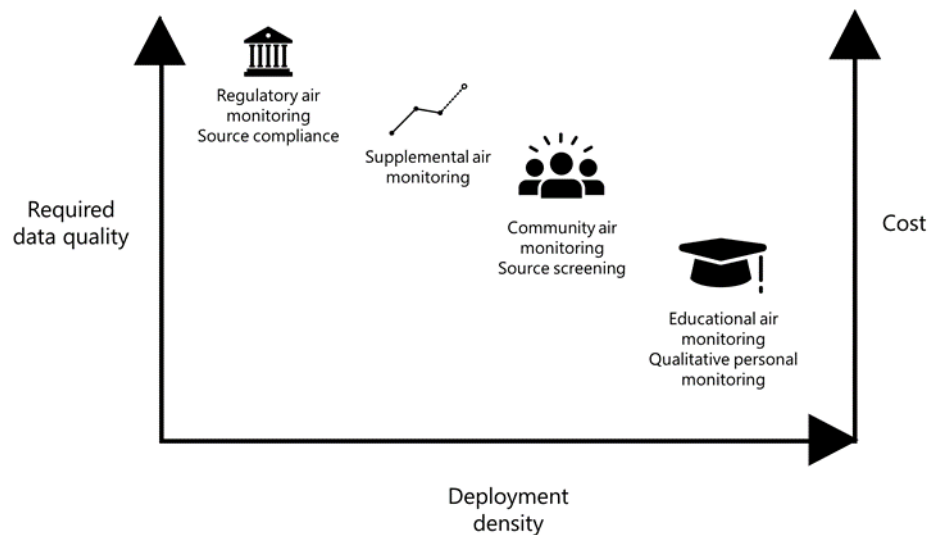


Figure 5. Trade-off between data quality, deployment density and cost (inspired by Snyder et al. 2013).

The literature review shed light on several key research areas that need to be explored in more detail (Figure 6), which can be sub-divided into three research fields (i.e. sensors, environment and actionable insights) as well as three additional topics at the intersection of these fields (i.e. sensor selection, calibration and correction). Major areas of interest in sensor performance assessment include sensor cross-responsiveness, repeatability in different environments and drift over time. With respect to impact of the environment, assessment of the impact of temperature, humidity and interferences need further assessment. Furthermore, opportunities remain to develop better methods for using sensor data to derive actionable insights (i.e. source control, exposure reduction and city planning). Given their large impact on sensor response, a comprehensive assessment of the impact of environmental factors is key for developing calibration protocols for low-cost AirSENCE devices. A detailed understanding of the impact of temperature and humidity as well as interferences on sensor performance is needed for development of better methods for understanding measurement uncertainty and correcting sensor response. More studies investigating the variation in sensor performance across a range of locations

with different pollutant levels and source profiles would be key for accomplishing this goal. Investigation of sensor performance across a range of sites would also help quantify the impact of calibration data on sensor response and will be key for designing better sensor maintenance protocols and calibration methods that lead to more generalizable calibration models. Regression techniques used for deriving calibration models could also benefit from more detailed assessment. In particular, a comparison of nonlinear and linear regression models as well as different variable transforms could help identify optimum methods for specific types and combinations of sensors. Comparison of regression methods for deriving cumulate health indices using low-cost sensor measurements could help develop more robust approaches for quantifying Air Quality Health Index. Further investigation of temporal sensor response patterns would help reveal new insights in level and characteristics of pollution sources, while also helping understand the causes for sensor drift and suggest methods that could be useful for correcting sensor response.

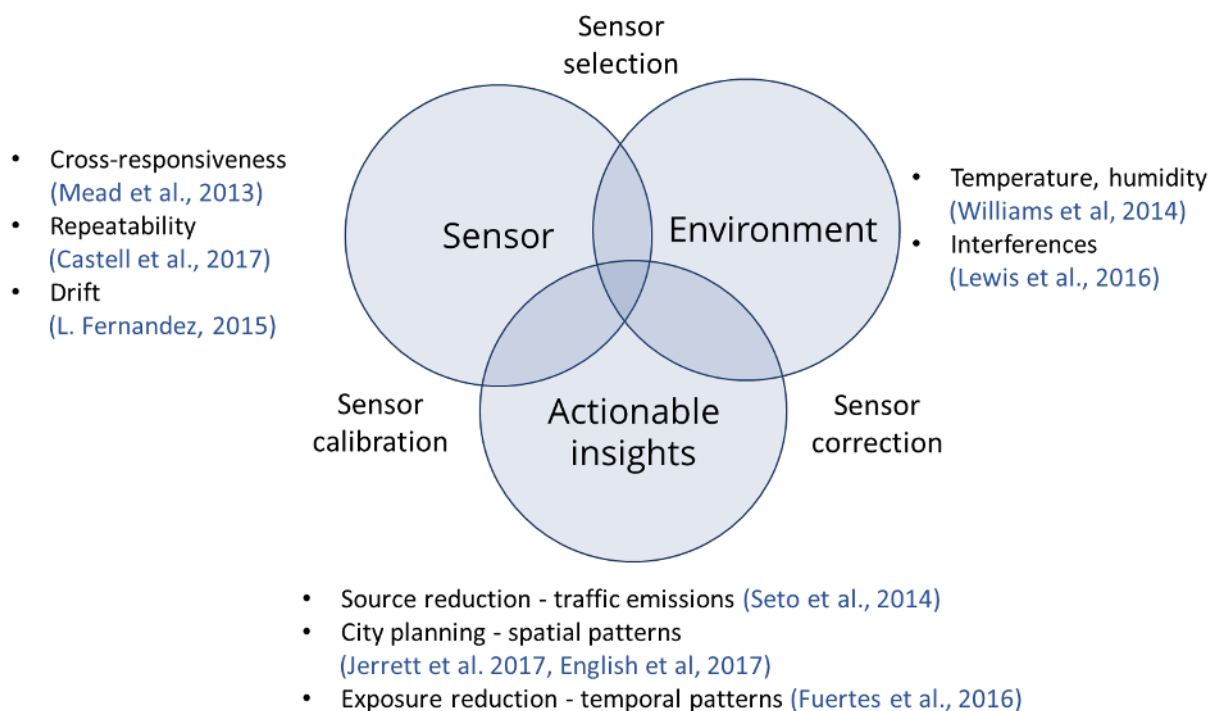


Figure 6. Overview of the three core areas in need for further research.

3 Methodology

The design and development of AirSENCE devices was conducted in modular manner and involved several key components. Sensor array printed circuit boards (PCBs) were designed in parallel with the data logging system and power supply PCBs and were integrated into a PCB assembly for each device. All the electrical components were contained in an enclosure which was designed to ensure optimum air sampling while reducing issues that are known to impact sensor response, such as the impact of humidity, overheating, contamination as well as uneven airflow dynamics. Once the devices were developed, they were calibrated and tested. Indirect calibration method was used, where the devices were calibrated by cross-comparison with reference equipment. Finally, the devices were validated and optimized in field during 3 sensor network campaigns carried out over the period of 5 years in Toronto.

3.1 AirSENCE device design

As highlighted in the literature review, very limited data is available on the evaluation of performance of pollutant calibration models as a function of sensor type or number (Section 2.1.4). A more comprehensive comparison of sensor performance across a range of commercial sensors is needed for designing an optimum sensor array for ambient air monitoring applications. Therefore, one of the goals of sensor device design was the setup of a reproducible platform for testing a wide array of different types of sensors prior to selecting a subset to be used in the models. Another key priority was evaluation of the sensor variability. To this end, several replicates of the device needed to be assembled to verify that the sensor response was comparable and correction factors were developed if needed. Finally, the analysis of different models of sensors was also important for understanding of sensor stability and sensitivity to interferences and drift. For example, as shown in the literature review, temperature and humidity are known to have a large impact on sensor performance and can contribute to drift (Section 2.3.1). The design of sensor enclosures needed to be such that the air sampling is consistent and representative of ambient air, while the impacts of temperature and humidity are minimized. Several devices were designed and tested over the duration of the project. Sensor array device design (Figure 7) consisted of the following four components: 1) enclosure and air circulation system, 2) data logging system, 3) power management system, and 4) sensor array system. These four components were developed and improved upon over the course of this PhD project. Several key target attributes were identified for each of the device components based on literature review and initial prototyping (Table 7).

The overview of system design is provided in Figure 7 and detailed design overview is provided in Appendix .

Table 7. Key attributes of AirSENCE device deign.

System component	Target attributes
Enclosure and air circulation system	<ul style="list-style-type: none"> • Fast air exchange • Consistent air sampling across all sensors • Minimized influence of weather, temperature and humidity
Data logging system	<ul style="list-style-type: none"> • Modular design for easy exchange of components • Removable data storage system • Flexibility in integrating additional capabilities (i.e. RF communication)
Power management system	<ul style="list-style-type: none"> • Adjustable voltage output between 2.5 and 8V required by the sensors • Stable power supply with low noise level • Modular design for easy replacement
Sensor array system	<ul style="list-style-type: none"> • Consistent sampling across sensors • Flexible connector system for easy exchange of sensors • Minimized thermal noise

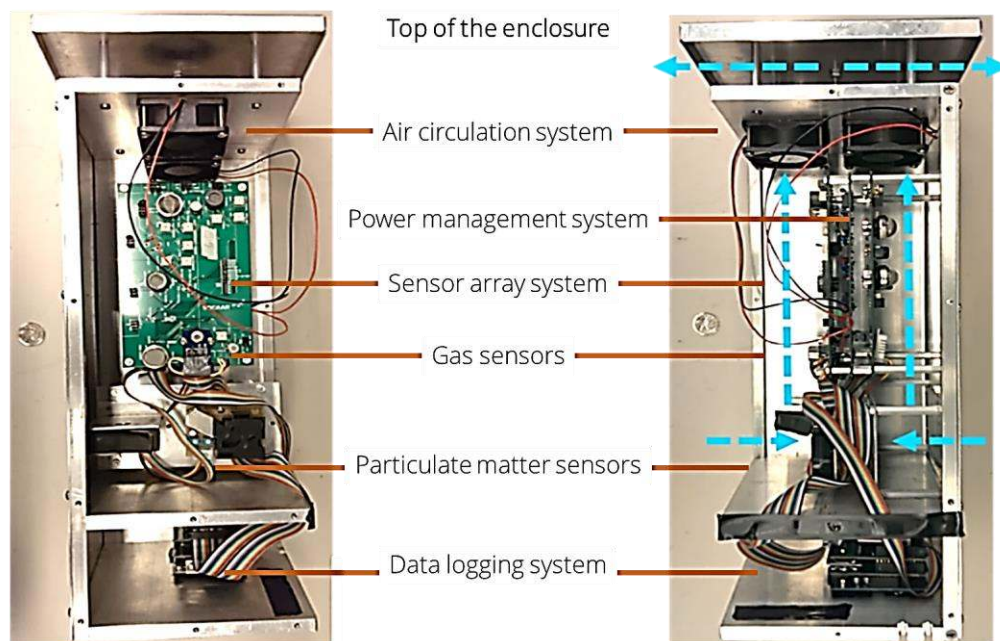


Figure 7. Diagram of AirSENCE device design.

3.1.1 Enclosure and air circulation system design

Over the course of the thesis project, several versions of AirSENCE devices were assembled. They were tested in different micro-environments such as indoor, outdoor and public transportation. The testing results provided insight into some of the challenges and tradeoffs for reliable monitoring air pollution in field. For example, exposing sensors to direct airflow improves the ability to detect high frequency variation in pollution levels but often results in a reduced sensor lifespan due to direct exposure to humidity and particulate matter suspended in the air. Other challenges included uneven airflow distribution, thermal noise and interferences. An overview of the challenges and design process is provided below, and a detailed overview of the design is provided in Appendix 2: Design for the AirSENCE sensor device.

3.1.1.1 Airflow dynamics and sensor distribution

Active air exchange is one of the most common type of air sampling in air quality monitoring field (Arfire, Marjovi, and Martinoli 2016). Active air sampling has benefits and drawbacks. While fans consume significant amount of power, they can reduce residence time and therefore increase sensor device sensitivity for detecting high frequency variation in air pollution. It is important that airflow through the sensor array is fast but sufficiently mixed to ensure that the sampled air mass is homogenous and representative of the ambient air. Therefore, a fan was used in AirSENCE devices for air exchange. In addition, the sensor array enclosure was designed to optimize performance. A vertical tube-like design was chosen to ensure laminar airflow and even air distribution while minimizing the impact of humidity and the contamination of gas sensors with residual particulate matter in the air. In this tube-like arrangement, air was sampled from the lower part of the device, pulled through the sensing chamber and exhausted in the upper part of the device (Figure 7). The air exchange inside the device was generated by two low-power fans in the top section of the enclosure. Even when the fans were turned off, low level of air exchange was still maintained as a result of convection generated by the heat gradient created by catalytic heating modules built in gas sensors. Airflow considerations have also informed the sensor placement for particulate matter sensors. Unlike gas sensors, they needed to be placed close to inlet to minimize particle losses due to impaction on the walls of the enclosure. In addition, particulate matter sensors were required to have a very specific orientation and position within the device to ensure continuous air flow through the optical chamber (Figure 7). Therefore,

special attachment points were designed for them in the lower part of the device directly in front of the inlet.

3.1.1.2 Overheating

Metal oxide gas sensors have built-in catalytic heating modules, which need to be operated at 200-300°C. Due to the large array of sensors tested, thermal noise due to the overheating of power lines was a concern. To protect microcontroller and sensitive ICs from overheating, while maintaining compact design for portability, a 2-layer container system was designed. A vertical tube-like enclosure allowed us to use the sensor heating to our advantage, creating convection as an additional mechanism for air flow through the device. The sensitive integrated circuits and microcontroller were physically separated from the high-temperature components and sensors using an additional compartment in the lower part of the device. The locations of sensors on the sensor PCBs was also chosen such that the thermal distribution and power consumption is balanced as much as possible to avoid high current in the power lines and thermal noise in signal paths. In addition, temperature and relative humidity sensors needed to be positioned further away from the heated MOS sensors so that they could be exposed to the same air mass but be protected from heating.

3.1.1.3 Interferences

The impact of dust, interferences and contaminants was another key challenge that needed to be addressed when designing the sensor enclosure. A large amount of dust accumulated inside the device when it was deployed for long periods of time outdoors. To reduce the impact of dust and insects, a mesh screen was set up on the inlet of the device. This screen could also be optionally converted to an impactor assembly if removal of larger particles was needed. Sensors were distributed on the PCBs such that the sensors for monitoring pollutants with similar redox chemistry are close to each other. This setup reduced the impact of pollutant depletion processes resulting from gas catalytic reactions (Figure 3) and ensured the data was consistent.

The stability of enclosure material and the potential interactions between the surface of the material and sampled air mass was another challenge. Although plastic was the preferred material for prototype development due to light weight and durability, there were some concerns that it could be off-gassing volatile organic compounds at high temperature generated by the gas sensors and interfering with the

sensor response. To address these issues, customized aluminum enclosure was designed. As a result, no additional waterproofing was needed. During the SCULPT 2013 summer campaign, the devices were placed on a tripod, allowing sampling at controlled height, reducing the contamination with dust (Figure 8).



Figure 8. Sensor array device setup during the SCULPT 2013 summer campaign.

3.1.2 *Electrical design*

Modular design was used for AirSENCE electrical system to facilitate troubleshooting and minimize the cost of fixing a device. Each device consisted of five printed circuit boards (PCBs): two sensor PCBs (each with a unique sensor subset), one power management PCB (interface board), one PCB for data logging and one PCB for signal multiplexing (Figure 9).

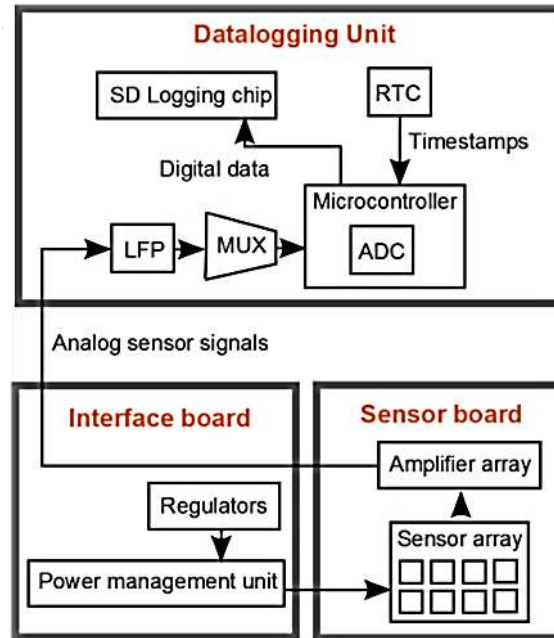


Figure 9. System level diagram of AirSENCE device design.

Several versions of prototypes were hand-wired before shifting to PCB design and fabrication. The first PCB-fabricated prototype was designed and assembled for the SCULPT study in August 2013 (AirSENCE Rev1). PCB manufacturing was outsourced, while component soldering and assembly was completed at the lab. Each device required five boards to be soldered, tested and assembled. An updated version of the device (AirSENCE Rev2) was manufactured in April 2014 in preparation for Pan Am study, while another major revision of the device (AirSENCE Rev2.5) was designed and manufactured in March 2016 in preparation for deployments in Beijing. In total, over thirty-five devices were manufactured over the course of the project (Figure 10). Over the course of the project, several revisions were made to device design (Table 8). Most of the analysis in this document is focused on AirSENCE Rev1-2.5 due to consistent sensor subset used in these device revisions as well as large number of data samples collected.

Table 8. AirSENCE device revision description.

Device version	Revision description
AirSENCE Rev0	<ul style="list-style-type: none"> • Device was hand-wired. • Sensor array was composed of 5 gas sensors (MQ3, TGS822, TGS825, MQ135 and TGS880) and 2 particulate matter sensors (ionization and optical sensors from smoke alarms).

	<ul style="list-style-type: none"> • Data logging through USB connection to a laptop. • Plastic Tupperware enclosure was used.
AirSENCE Rev0.5	<ul style="list-style-type: none"> • Device was hand-wired. • Sensor array was composed of 5 gas sensors (MQ3, TGS822, TGS825, MQ135 and TGS880), DHT22 temperature and humidity sensor and 2 particulate matter sensors (GP2Y1010AU0F, PPD42). Particulate matter sensors were upgraded because smoke detector sensors did not show good performance. Temperature and humidity sensor was added to improve the accuracy of regression models. • Internal data logging module was integrated in the device to enable data collection without tethering to a laptop. • Enclosure was upgraded to a more robust design and power adaptor was used to power the device.
AirSENCE Rev1	<ul style="list-style-type: none"> • Circuit schematics were transitioned to PCB design. The printing was outsourced but the assembly was done at the lab. • Sensor array was composed of 12 gas sensors (MQ3, TGS822, TGS825, MQ135 and TGS880, TGS2600, TGS2620, MICS5525, MICS5521, MICS5121, MICS2611, P/N706), DHT22 temperature and humidity sensor and 2 particulate matter sensors (GP2Y1010AU0F, PPD42).
AirSENCE Rev2	<ul style="list-style-type: none"> • Several major circuit design errors were fixed in an updated PCB design. • Wireless capability was added to the device.
AirSENCE Rev2.5	<ul style="list-style-type: none"> • Power management interface board was revised to accommodate for more efficient power regulation. • Website was developed to display sensor data.
AirSENCE Rev3	<ul style="list-style-type: none"> • Two sensor boards were consolidated into one board. • The list of sensors was overhauled due to sensor model obsolescence. • Device design was reduced in size and weight. • Device design was made more weatherproof.

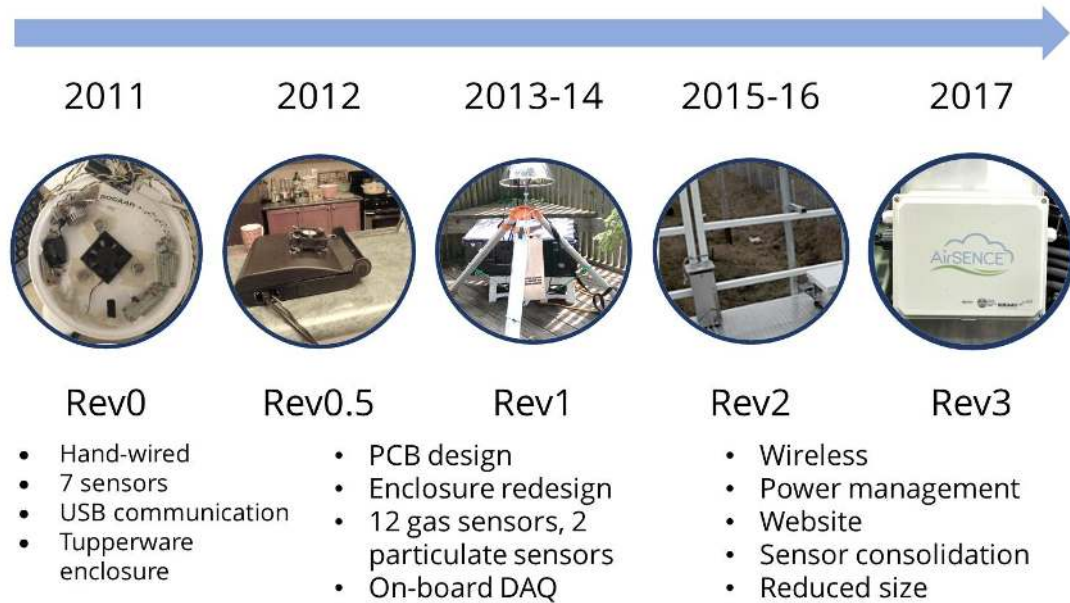


Figure 10. AirSENCE device design iterations.

3.1.2.1 Data acquisition system

Data acquisition system was designed for signal sampling, time keeping and data storage. The key components of the system included a microcontroller module and a data logging system. A commercial microcontroller module (Arduino) was used for controlling the operation of the sensor device and providing access to peripheral inputs and outputs, such as USB and DC power supply. It included a signal processing circuit and a microcontroller circuit. The module also included power regulation circuit and adaptors for communications (USB port) and power source (AC/DC adaptor). Arduino microcontroller module was used as a backbone for the data acquisition system because it provided a balance between price and functionality. Arduino Uno Rev3 microcontroller (~\$25 USD) was used for most versions of the device and included 13 digital input/output pins, 6 analog input/output pins and a built-in 10-bit analog to digital converter (ADC). 10-bit resolution was found to be sufficient for the sensors we used. Data logging system, assembled in the lab, included low pass filter, real time clock (RTC), SD card flash storage module, and multiplexing module. Real time clock (RTC) IC was used for generating time stamps. SD card

flash storage module was used for data storage. Low pass filter was used for circuit noise reduction, while 16:1 multiplexer was needed to expand the number of input pins for testing more sensors.

3.1.2.2 Power management system

Careful design of voltage regulation system was important to ensure sensors were operated at correct and consistent voltage while also maintaining the device power efficiency. Many of the tested sensors had different requirements for supply voltage (ranging between 1.7 and 8V) and current (ranging between 10mA and 1A). As mentioned earlier, metal oxide gas sensors have built-in catalytic heating modules, resulting in high current draw. Switching regulators were therefore selected to produce stable output voltage while ensuring energy efficiency. Operating voltage of the system was set to 5V and five regulators were used to control the voltage - four buck converters were used to down-regulate the voltage and one boost regulator was used to up-regulate the voltage for sensors and components requiring lower and higher voltages respectively. Power management PCB was also used for managing the operation and power supply to external components, including the fans controlling device airflow and Arduino microcontroller. In addition, power management PCB was used to interface AirSENCE device components to simplify the design and to reduce the number of cables needed. Sensor signal traces were also passed from the sensor boards to data logging module through the power management PCB.

Device power consumption was evaluated and different scenarios for power consumption reduction were tested. Original power consumption of the device was on the order of 7W (5V, 1.4A) for the temperature range of 10-22 °C. Sensor subset optimization described in Section 4.1.5 allowed for reduction in power consumption to 5.25W (5V, 1.05A). Tests with commercial 20 Ahr battery packs showed that the device could be operated continuously for 6 hours, which could be significantly increased if several battery backs were combined.

3.1.2.3 Sensor system

The sensor system consisted of two sensor PCBs (each with a unique sensor subset) attached to interface board through a PCB-mounded connector. Both sensor PCBs were modular and could be easily replaced in case of failure. The sensors mounded on each of the PCBs could also be easily removed and replaced at the end of their operational lifetime. Sensor PCBs have undergone three major revisions

over the course of the project, although many smaller sensor subset revisions were also implemented whenever new sensors were released commercially, or sensors were replaced due to poor performance.

Several different gas sensor types were considered for gas pollutant monitoring using AirSENCE, including metal oxide semiconductor sensors, electrochemical sensors, photoionization sensors and IR sensors. Metal oxide semiconductor (MOS) gas sensors were selected, due to low cost, small size, simple signal monitoring circuit and availability of many different sensors for a range of gas contaminants. Gas sensor output was measured using an adjustable voltage divider, created using a potentiometer. Sensor resistance was calculated using sensor input voltage, load resistance and voltage across the load resistor in the voltage divider circuit. The use of the potentiometer to adjust the value of load resistor allowed for the control of output voltage. Due to low manufacturing reproducibility associated with low-cost sensors (Section 4.1.1), the ability to adjust sensor output was important.

Several different types of particulate matter sensors were tested, including ionization-based particle sensors and photoelectric particle sensors extracted from smoke detectors, as well as optical light scattering particulate matter sensors sold as components commercially. Light scattering particulate matter particle sensors were selected because of low cost, ease of use and promising results cited in the research literature (Section 2.1.2). The particulate matter monitoring circuit was more complicated than the gas sensor circuit, requiring drivers and amplifiers to control the operation of the light source LED and process the output of the photodetector. However, most of the required circuit components were included in the commercially-available sensor circuit, making it possible to connect the sensor directly to the data acquisition system.

3.2 Calibration, data acquisition and data processing

Calibration process plays a key role in sensor performance in monitoring ambient air quality. As discussed in Section 2.1, the low manufacturing reproducibility, cross-sensitivity and nonlinearity of low-cost commercial sensors makes calibration an essential part of sensor device development process. While some calibration information is available in sensor datasheets, it is limited to lab conditions and is often misleading. While direct lab calibrations using diluted pollutant standards are typical for air quality monitoring equipment (Ministry of the Environment and Climate Change 2017), these methods were shown to have inadequate performance for the low-cost sensors due to the high level of cross-sensitivity as well as sensitivity to environmental interferences (Section 2.2.1). Pollution composition is

highly spatially variable in urban environment, which is a challenge for cross-responsive sensors. Therefore, variability of the training dataset used for development of calibration models is key for ensuring good performance. As a result, one of the key areas of focus of this thesis was the development of robust calibration protocols optimized for low-cost sensors. Another key area of focus was the design of an optimum data analysis pipeline. It was important to develop protocols for correcting sensor response for outliers, noise, nonlinearity, inter-device variation and drift. These factors have a large impact on accuracy and precision of sensor response.

3.2.1 Calibration process

As indicated in the literature review, calibration under lab conditions generally does not result in calibration models that perform well in ambient environment (Section 2.2.1). Therefore, indirect calibration approach was used in this work, where AirSENCE devices were co-located with standard air pollution monitoring equipment for a period of time to generate calibration datasets. AirSENCE devices were co-located directly outside of the SOCAAR lab window at 200 College Street, sampling air from the College St and St George St intersection (Figure 11). Sensor calibration was conducted by comparing sensor response to standard instrument response over a period of 1 week. Five standard instruments were used for calibration – Thermo 42i NO_x monitor, Thermo 49C Ozone monitor, Thermo 48C CO monitor, Thermo 410i CO₂ monitor and Magee Scientific Sharp AE33-7 (continuous light scattering nephelometer). In addition to pollutant calibration, AirSENCE devices were also calibrated against cumulative multi-pollutant index air quality health index (AQHI).

AirSENCE devices were set up to sample air using fan-driven flow-through system. The reference instruments were sampling air 10-15 m west from the sensors, along the same side of the building. Indirect calibration also has several limitations. Because of the co-presence of many different pollutants in ambient air, sensor selectivity cannot be directly evaluated. In addition, pollutant concentration and environmental interferences cannot be controlled, resulting in unreliable calibration model performance at high pollutant levels. Several approaches to expand the calibration data variability were evaluated, including the use of historical calibration data and calibration of devices at several locations with diverse pollutant levels and source profiles. The results of these evaluations are discussed in Section 4.4.4.



Figure 11. Sensor array device calibration setup.

3.2.2 Data acquisition and processing

As mentioned above, development of protocols for correcting sensor response is key for ensuring reliable device performance due to the significant impact of an array of factors on sensor performance. These factors include outliers, noise, nonlinearity, cross-sensitivity, inter-device variation and drift. Data acquisition and processing software pipeline was developed, composed of: 1) data logging C++ firmware running on the device microcontroller, 2) data reformatting, averaging, merging and cleaning Python script, and 3) baseline correction and outlier removal R script, and 4) data exploration, variable selection and model development R scripts. (Figure 12).

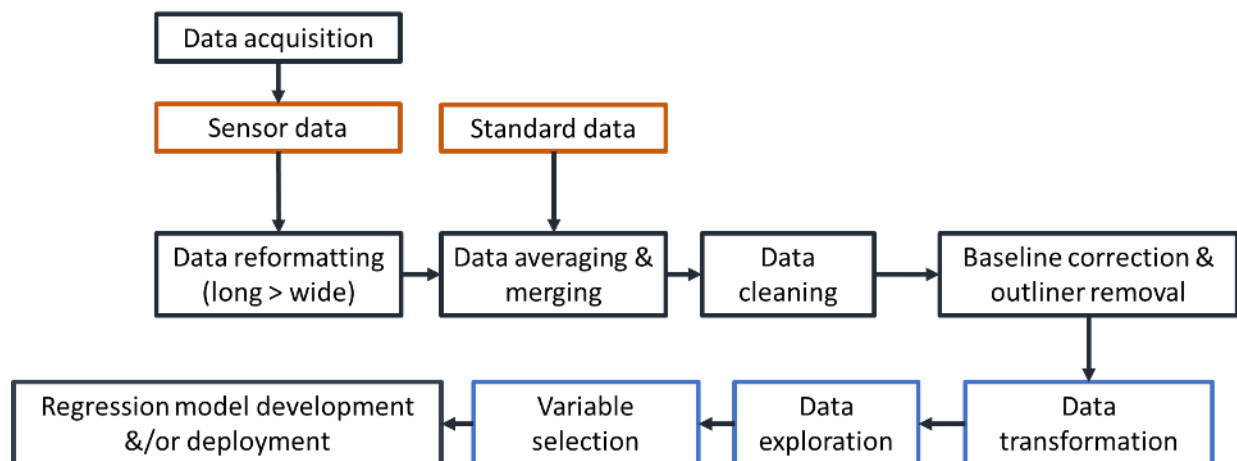


Figure 12. Data processing pipeline. Black colour denotes steps that are always part of the analysis, blue colour denotes steps that are only relevant to some type of analyses. Orange colour denotes data sources.

Data acquisition process describes the steps involved in extracting data from device. Microcontroller firmware was developed to collect and log time stamps and sensor measurements while ensuring correct operation and sampling frequencies for each of the sensors. Initially, Arduino platform was used, which was based on ATmega328P microcontroller and runs machine code compiled from C++. While libraries were available for some modules (i.e. Real Time Clock, SD card logging), other libraries needed to be developed (i.e. sensor data probing and signal processing). Initially, the firmware functionality was limited to signal processing and data logging on an SD card. In the later versions, the ability to communicate using Xbee, WiFi and GPRS protocols was added. In newer versions of the device, the data was wirelessly communicated between devices (i.e. Xbee protocol) or sent to an online server (WiFi and GPRS protocols). This was coupled with software for reading data from the online server and backing it up in the local MS SQL database. As functionality was expanded, the built-in Arduino module memory proved insufficient for the complexity of procedures. Therefore, Arduino platform was substituted for standalone ATmega microcontroller with more available memory. This transition required most of the code and many of the libraries to be re-written but allowed for more flexibility and more complex processes to run on the microcontroller.

The next step of the data processing pipeline was data reformatting. Statistical computing language R was used for initial data and time stamp reformatting⁵. Data reformatting was necessary because the optimum format for data logging was different from the optimum format for analysis. Long format was used for data logging to capture high time resolution sensor data. However long format was not optimum for data review, cleaning and analysis.

Data reformatting was followed by data averaging and merging using Python. Data was averaged at 1, 5, 15, and 30min resolution depending on the type of analysis. A separate step was needed for data averaging and merging because time resolution and data continuity varied between different reference instruments.

The fourth step of data processing pipeline was data cleaning, which included purging data off flagged measurements (i.e. shifts in sensor readings, outliers) associated with sensor stabilisation, device

⁵ <http://www.r-project.org>

troubleshooting, and optimisation. Data cleaning was an important step in data processing for several reasons: device testing⁶, adjustments, and stabilisation. Sensor device testing was regularly conducted to identify and correct any performance issues early (i.e. faulty sensors, components, and circuit errors). Sensor operation range was also adjusted for the pollutant levels using the potentiometers, which resulted in a shift of baseline voltage level. Another important factor, sensor stabilisation, is defined here as decrease in sensor response noise due to heating of metal oxide semiconductor film. For new sensors, stabilisation time was 24 hours. For sensors, which have been previously used and were later disconnected for sufficient time to allow the MOS sensor to cool down (i.e. greater than 1 minute), stabilisation time could range between 15 minutes and 1 hour. Sensor response during stabilisation had different characteristics compared to sensor response post-stabilization due to adsorption and desorption of gases at the MOS surface. As a result, the measurements associated with device testing, adjustment and stabilisation were flagged and removed.

Another important correction required for all sensors was baseline correction. In particular, MOS sensors have a large baseline resistance, which varies significantly between replicate sensors of the same model. Therefore, baseline correction is an essential step in data preprocessing for these types of sensors. Baseline correction was performed by rescaling the data between 0 and 1 using R: the global minimum⁷ was subtracted from each observation and the result was divided by the difference between the global maximum⁸ and the global minimum. To avoid rescaling range bias, outliers were removed from sensor response by applying 1-99% bandpass filter. For some analyses, nonlinear transforms were applied to data to avoid the impact of sensor nonlinearity on calibration model performance.

The last 3 steps of the data analysis pipeline (i.e. data exploration, variable selection and regression model development) are more complex and vary depending on the type of analysis conducted. They are therefore discussed at length in Sections 4.1, 4.2 and 4.4.

3.3 Field study procedures

⁶ Device testing included: 1) basic PCB design and continuity tests, 2) power tests, 3) sensor circuit performance tests and optimisation, 4) data acquisition system tests, 5) wireless communication tests. For newly-assembled devices or those with sensors replaced, the initial week of sensor measurements was used for device testing.

⁷ Global minimum = the minimum sensor reading over one week of measurements after sensor stabilization.

⁸ Global maximum = the maximum sensor reading over one week of measurements after sensor stabilization

There were several key objectives for conducting field studies. The first objective was validation of AirSENCE device performance. As mentioned above, indirect calibration method had several limitations. One of them is the limitation in ability to evaluate the impact of interferences on sensor response. In addition, pollutant concentration and environmental interferences vary across sites, which could result in degraded calibration model accuracy at sites other than the calibration site. Therefore, validation of the device performance at several locations with diverse pollutant levels and source profiles was needed. This was accomplished by deploying AirSENCE devices at existing government pollution monitoring sites. Such validation studies have also allowed for evaluation of the impact of calibration training data variability on sensor performance. Another key objective for conducting field studies was the assessment of spatial variation of air pollution and the extent to which it can be quantified by AirSENCE devices. Deployment of AirSENCE devices at a range of different sites allowed comparison of pollutant models in terms of their ability to adjust for interferences and the extent to which their performance is generalizable. The third key objective of conducting field studies was evaluation of the ability of AirSENCE devices to resolve pollution sources. Different sources are associated with different pollution profiles and hence one of the key objectives was to test the ability of AirSENCE devices to detect influence of sources and feasibility of ranking sites based of the local pollution and source contribution. In addition, temporal variation in air pollution consists of several different components that can be attributed to different types of sources. Therefore, the feasibility of using high frequency variation in sensor response as a proxy of local traffic emissions was also of interest. The fourth key objective of field studies was design of better study protocols. For example, factors such as wind direction, time of the day and day of the week are known to have an impact on pollution levels. However, the impact of these factors on ability of AirSENCE devices to resolve sites has not yet been tested. Field studies offer the opportunity to develop better study protocols for evaluating spatial variation of air pollution as well as correcting for factors negatively impacting sensor response, such as drift. Four different campaigns were conducted to address the objectives describes above (Figure 13):

- Spatial Characteristics of Ultrafine Particles in Toronto (SCULPT) conducted in July-August 2013
- Long-term performance study conducted August 15, 2013 - May 10, 2014
- Pan Am study conducted in May-September 2015
- Multi-site validation study conducted between December 2015 and February 2016 at SOCAAR



Figure 13. AirSENCE device testing campaigns.

Four datasets collected over the course of the campaigns are summarized in Table 9.

Table 9. An overview of the data sets collected using AirSENCE devices during field studies.

	Dataset #1	Dataset #2	Dataset #3	Dataset #4
Campaign/ Study	SCULPT	Long-term performance study	Pan Am Campaign	Reference validation
Device version	AirSENCE Rev1	AirSENCE Rev1	AirSENCE Rev2	AirSENCE Rev2
Locations	6	1	7 (4)	3
Period	August 8 - 25, 2013	August 15, 2013 - May 10, 2014	April 20 - September 9, 2015	September 20, 2015 - January 15, 2016
Size	818 observations	6561 observations	191812 observations	139742 observations
Sampling resolution	30-minute resolution	30-minute resolution	1-minute resolution	1-minute resolution
Retention	90% of data used	70% of data used	90% of data used	90% of data used

3.3.1 *Spatial Characteristics of Ultrafine Particles in Toronto (SCULPT) Campaign and long-term performance study*

AirSENCE devices were deployed at 18 different sites in Toronto for two two-week deployments in July-August 2013 (Figure 14). In addition to sensors, reference instruments (CPC, DiscMini, SMPS, microAeth) for monitoring ultrafine particulate matter (UFP) and black carbon (BC) were also deployed at the sites. Eleven sites were investigated during the first deployment and seven sites were investigated during the second deployment. AirSENCE devices were moved between the central site and two remote sites, spending the first two weeks at one of the eleven sites (deployment 1), followed by one week at the SOCAAR lab, followed by two weeks at one of seven other sites (deployment 2). Measurements taken at the SOCAAR lab (August 5 – August 14, 2013) were used for sensor calibration. This gave an opportunity to test sensor response under a variety of different conditions. The spatial variability analysis is focused on 6 sites assessed during deployment 2 due to issues with device reliability during earlier deployments. As a follow-up to the SCULPT Campaign, long-term performance study was conducted at SOCAAR site to evaluate the performance of AirSENCE devices over longer period (August 15, 2013 - May 10, 2014) to assess drift and the impact of seasonal factors.

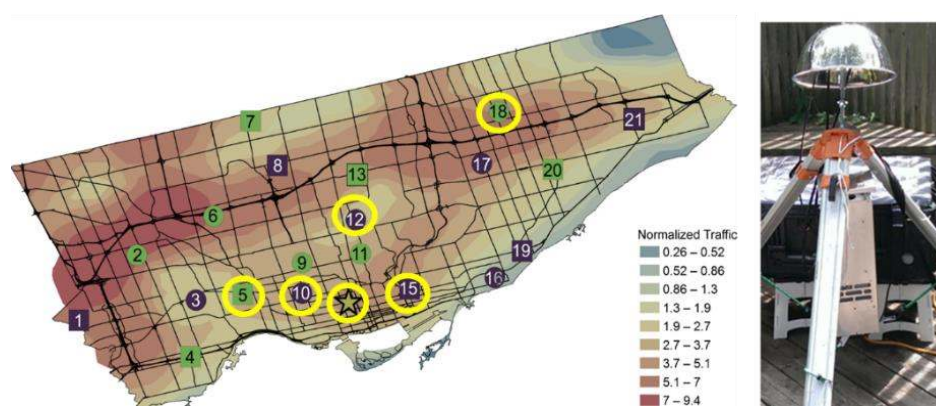


Figure 14. Selected sensor array device locations during SCULPT deployment. Devices from deployment 2 (outlined by yellow circles) were selected for analysis of spatial variation of pollution. Device setup during deployment is also shown in the photo from one of the sites.

3.3.2 *Pan Am study*

A summer campaign conducted during 2015 Pan American games in Toronto gave an opportunity to further test sensor response under a variety of different conditions. An updated version of the sensor device (AirSENCE Rev2) was tested in May-September 2015 in a multi-site deployment. AirSENCE devices were calibrated at SOCAAR for one week and deployed at eighteen different sites for periods of time ranging from seven days to one month. Four sites were selected for more detailed analysis (Figure 15). While no reference data was available at most sites, 2 locations (SOCAAR lab and MOE Station at Highway 401 site) had co-located standard instruments for NO_x, O₃, CO, CO₂, and PM_{2.5}.



Figure 15. Selected sensor array device locations during Summer 2015 Pan Am Games Campaign. Photos of device setup during deployment are also shown.

3.3.3 Multi-site validation study

Another multi-site study was conducted between December 2015 and February 2016 at three sites with standard instruments for NO_x, O₃, CO, and PM_{2.5}. The three sites included SOCAAR, Halan's Point at the Toronto Island and MOE station at Highway 401 (Figure 16). Two devices were deployed at each of these sites after calibration for at least a week at SOCAAR. The presence of reference pollution monitoring equipment at these sites allowed us to validate the sensor device performance as well as evaluate the impact of factors like temperature, humidity, wind direction and drift.



Figure 16. Sensor array device locations during Winter 2015 validation at Highway 401, SOCAAR and Halans Point. Photos of device setup during deployment are also shown.

4 Results and discussion

4.1 Sensor performance

The first major aim of this research work was to evaluate commercially-available low-cost sensors with regards to their performance and identify the sensors that are most suitable for applications in long-term field deployments. A secondary goal was to identify methods which could be used for correcting sensor performance issues. The following research questions were of interest:

Sensor response variation

1. What is the extent of inter-device variation for the commercially available sensors?
2. What is the distribution of the inter-device variation? How does it vary with sensor response level?
3. What is the effect of environmental factors and pollutant levels on the variation between devices?

Sensor importance, distribution and multicollinearity

4. What are characteristics of sensor data distribution and how does it compare to air pollutant distribution?
5. What is the effect of multicollinearity on sensor response and which sensors are the most affected?
6. Which pollutants can be most readily measured by commercially available AirSENCE devices?

Site influence & drift

7. What is the extent of correlation between sensor response and air pollutant levels? How does this change when sampling at a different location (i.e. a different pollutant mixture)?
8. What is the extent of drift in sensor baseline and range and which sensors are affected the most?

4.1.1 *Intra-sensor variation is a big challenge*

The initial goal for sensor evaluation was to quantify the extent of variation between replicates of the same sensor. This evaluation is important to understand the degree of error in sensor readings and is especially important for air pollution monitoring sensor networks. Accurate quantification of spatial pollution patterns depends on the assumption that the extent of variation between sensor replicates is

minimum. Knowledge of the inter-device sensor variation also allows us to assess the extent to which calibration factors can be transferred (i.e. calibration factors and regression coefficients derived from one device data can be applied to other devices). In addition, most sensor datasheets do not quantify the extent of variation between sensors and comprehensive evaluation of this factor has not yet been conducted in the research literature.

Five replicates of 15 different commercial sensors were selected for detailed evaluation (Table 10). The sensors were selected from a much larger pool of evaluated sensor models (Appendix) based on a combination of factors, including sensor cost, initial performance, circuit complexity, size and power consumption. To evaluate the sensor performance, 5 identical devices were assembled, each containing a copy of each of the sensors shown in Table 10. Sensors were co-located on an outdoor windowsill at SOCAAR and their performance was evaluated 4-8 months. Inter-device sensor variation was determined for each sensor, derived from standard deviation and mean of rescaled sensor measurements.

Table 10. Sensor subset included in each AirSENCE device.⁹

Manufacturer	Model	Operation	Price (\$CAN) for 1-9 units	Sensor detection targets identified in the sensor data sheets
Aosong Electronics	DHT22	Thermistor, capacitive sensor	\$15	Temperature, humidity
Figaro	TGS2600	MOS	\$15	Methane, CO, iso-butane, ethanol, hydrogen
Figaro	TGS2620	MOS	\$15	Methane, CO, iso-butane, ethanol, hydrogen
Figaro	TGS822	MOS	\$15	Organic solvents, acetone, methane, CO, isobutane
Figaro	TGS825	MOS	\$57	Hydrogen sulfide
Figaro	TGS880	MOS	\$15	CO, methane, isobutane, ethanol, cooking vapours

⁹ <https://www.yumpu.com/en/document/view/379764/price-delivery-information-for-figaro-gas-sensors>

E2V / MICS	MICS5525	MOS	\$10	CO
E2V / MICS	MICS5521	MOS	\$10	CO, hydrocarbons, volatile organics
E2V / MICS	MICS5121	MOS	\$10	CO, hydrocarbons, volatile organics (protective filter)
E2V / MICS	MICS2611	MOS	\$10	O ₃
Synkera	P/N706	MOS	\$30	NO _x
Hanwei	MQ3	MOS	\$5	Alcohol, benzene, CH ₄ , hexane, LPG, CO
Hanwei	MQ135	MOS	\$8	NH ₃ , NO _x , alcohol, benzene, smoke, CO ₂ , CO
Sharp	GP2Y1010 AU0F	Optical	\$12	Particulate matter
Shinyei	PPD42	Optical	\$12	Particulate matter

Overall, the variability between AirSENCE devices was found to be high, suggesting that sensor readings are not very reproducible (Figure 17). Most MOS gas sensors and optical PM sensors are characterised by high inter-device sensor variability. Sensors P/N706 (NO_x) and PPD42 (PM₁) were found to show the highest inter-device variation, while temperature and humidity sensors - the lowest. These findings suggest the importance of inter-device comparison and calibration prior to their deployment in field studies.

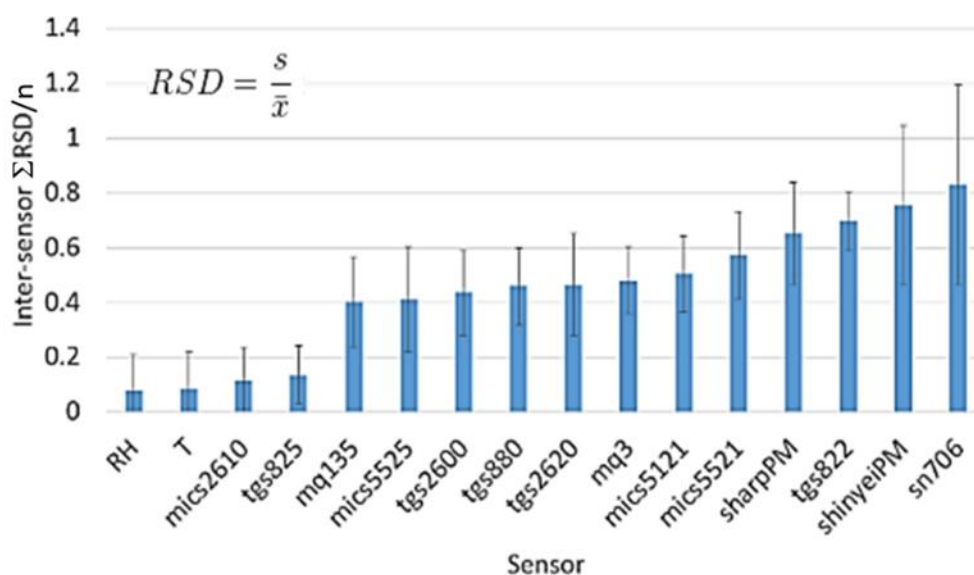


Figure 17. Inter-device sensor variation of sensor response. Five AirSENCE devices were compared based on raw voltage readings rescaled 0-1. %RSD across the five devices was derived for each 30-minute observation and the average %RSD along with the standard deviation of these average %RSD values are shown. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

4.1.2 *Variation in sensitivity*

As inter-device sensor variation analysis indicated, most sensors show high variability between replicates and need to be corrected. There are several methods commonly used for correcting for inter-device sensor variability. These include:

- Rescaling sensor reading
- Correcting sensor response based on multi-sensor average
- Deployment of redundant sensors within each device

To establish the type of correction required and whether a correction is possible, the distribution of inter-device variation has to be evaluated. The relationships between each of the sensors to an average of all sensors of specific type was found through correlation.

It was found that for some sensors, the relationship between sensor replicates is a simple bias that can be corrected by baseline removal through rescaling. For other sensors, the inter-device sensor variability function was more complicated. Transformation of sensor data through power or log transforms is often required to correct the data due to non-linearity at high sensor response levels (Chaiyaboun et al. 2007; Hirobayashi et al. 2003). However, for most sensors the inter-device variation is caused by the variation in slope between devices (Figure 18), which suggests that major reasons for large inter-device sensor variation is the difference in sensor sensitivity. For TGS822 gas sensor (a sensor for monitoring volatile organic compounds), the response of each sensor replicate was linear relative to sensor average, but the slope varied between 0.5 and 2.

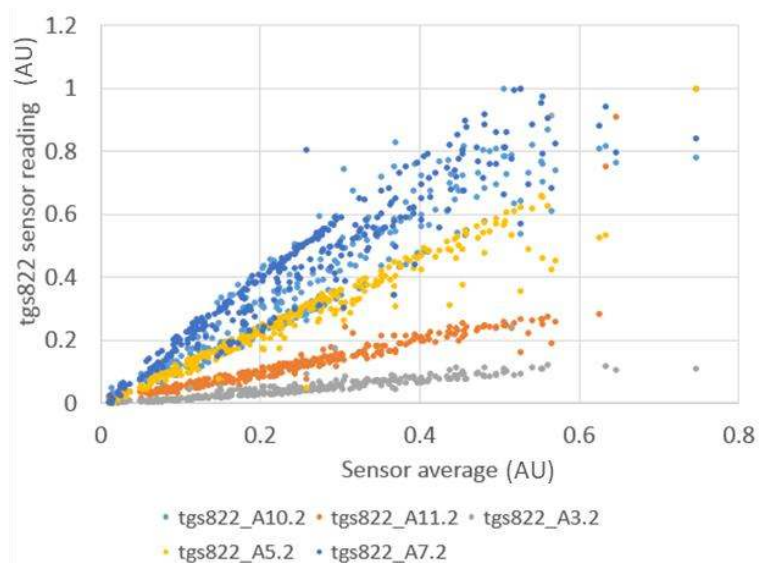


Figure 18. Inter-device sensor variation for the TGS822 sensor. TGS822 is a MOS sensor for volatile organic compounds, such as solvents. The units of the data shown are arbitrary (sensor voltage measurements rescaled 0-1). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

This type of variation can also be corrected for using a linear function (i.e. multiplying by a constant). An example of the impact of sensor response correction is shown in Figure 19. Sensor response was corrected by co-locating AirSENCE devices and deriving a correction function based on relationship between the sensor response and that of the average of the AirSENCE devices.

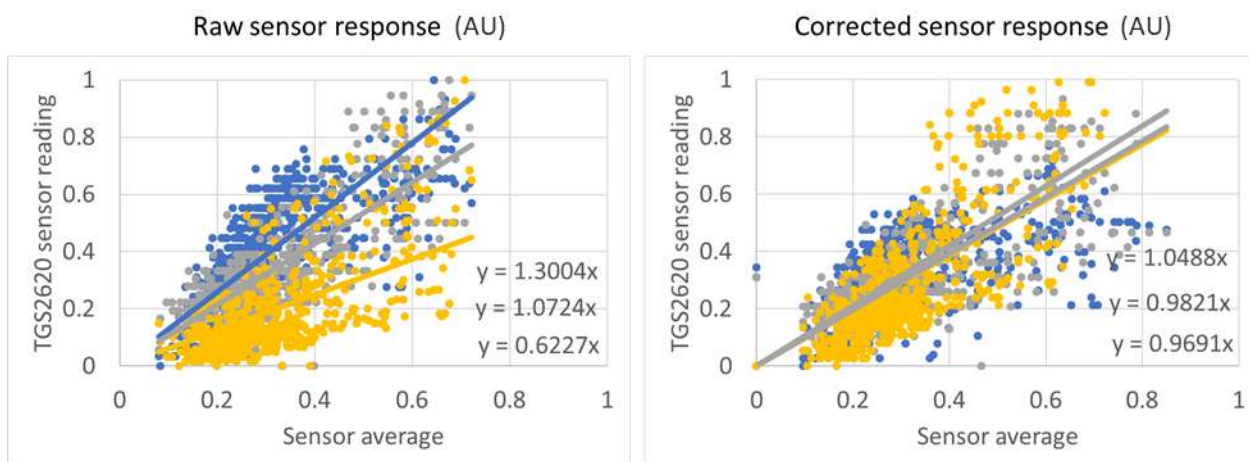


Figure 19. The impact of sensor response correction on reducing inter-device sensor variation. Three AirSENCE devices are compared. The units of the data shown are arbitrary (sensor voltage

measurements rescaled 0-1). Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

The high variability between sensor replicates as well as presence of different types of inter-device sensor variability contribute to systematic error and highlight the need for device-specific calibration. During device-specific calibration, intercepts and weights are derived separately for each device replicate during calibration, allowing to correct for both additive bias and multiplicative sensitivity variation.

4.1.3 *Influential factors*

While most of the instances of inter-device sensor variation were systematic in nature, sensor response was also influenced by interference error and random error, which are less consistent and more difficult to predict.

$$\text{Sensor output} = \text{Signal} + \text{Systematic error} + \text{Interference error} + \text{Random error}$$

For example, the variation in response of the TGS2620 (a volatile organic compound sensor) was affected by inter-device sensor variation that is not only due to variation between sensors but also the variability in impact of temperature on sensor response (Figure 20). Higher ambient temperature (above 15 C) was associated with higher variation between devices. MOS sensor response is catalyzed by increase in temperature generated by the integrated heating element. High impact of temperature on sensor response as well as high variability in temperature response profiles of MOS sensors has been shown in the literature (Masson, Piedrahita, and Hannigan 2015; Huerta et al. 2016). Hence inter-sensor variation in the temperature-sensor response profile could be one of the causes for the higher spread of the sensor measurements at higher temperatures. These results suggest that sensor-specific calibration under a range of ambient temperatures is essential for accurate results.

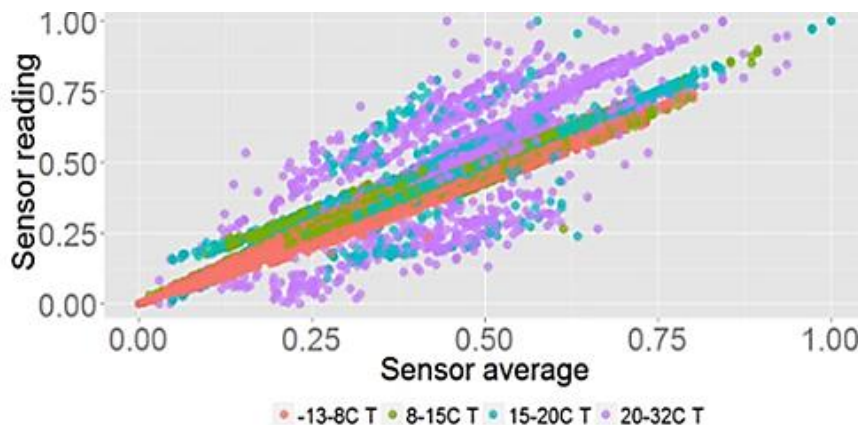


Figure 20. Inter-device variation of tgs2620 sensor and the impact of temperature. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

The impact of environmental factors on sensor data was systematically evaluated by binning the sensor measurements using the 4 quartile range bins of different factors (i.e. environmental variables and pollutant concentrations). The difference between 1st and 4th quartile was used as metric of the influence of the factor.

The influential factor analysis indicated that temperature, O₃ and CO concentrations were the biggest contributors to high variation between devices (Figure 21). High temperature had greater influence on inter-device sensor variability than low temperature, while higher pollutant concentration always contributed to higher variation between sensors. The large influence of temperature is expected due to the importance of temperature in the operation of MOS gas sensors (i.e. the catalytic effect of temperature on sensor response).

One of the major reasons for increase in inter-device sensor variation with increase of concentration is the nonlinearity in sensor response and high inter-sensor variability in sensitivity at higher pollutant levels. The variability between devices is smaller at low pollutant levels because the change in sensor response with respect to change in pollutant levels is smaller and more consistent. The challenges associated with variation in sensor sensitivity as well as opportunities for sensor response correction are discussed in more detail in sections 4.2.4 and 4.4 respectively.

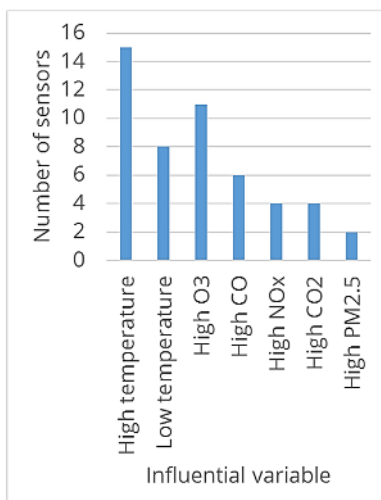


Figure 21. Summary of variables contributing to increasing inter-device %RSD. For example, 15 of the 16 sensors were influenced by temperature. Only consistent trends that apply across all sensor replicates are shown.

4.1.4 Sensor multicollinearity and distribution

Many of the commercial gas sensors are cross-responsive (Masson, Piedrahita, and Hannigan 2015). However, this lack of specificity creates an opportunity for multi-sensor monitoring and analysis using pattern recognition. Therefore, analysis of the correlation between sensors and pollutant monitors is important to understand sensor response characteristics and identify potential issues with predictor collinearity for regression models utilizing multiple different types of sensors. Even though multicollinearity does not reduce the overall predictive power or reliability of a model, it does affect predictor regression coefficients and relative importance analysis. In addition, this analysis helps identify sensors which could be used for correcting the model output. For instance, one of the sensors might be very sensitive to CO but is affected by drift while another sensor might be less sensitive to CO but is also less affected by drift. Including both sensors in the model would introduce redundancy but could lead to a calibration model that can accurately measure CO while being less prone to drift. To establish the degree of correlation between sensors as well as correlation between sensors and reference monitors, initial data exploration through evaluation of correlation matrices was performed (Figure 22, Figure 23).

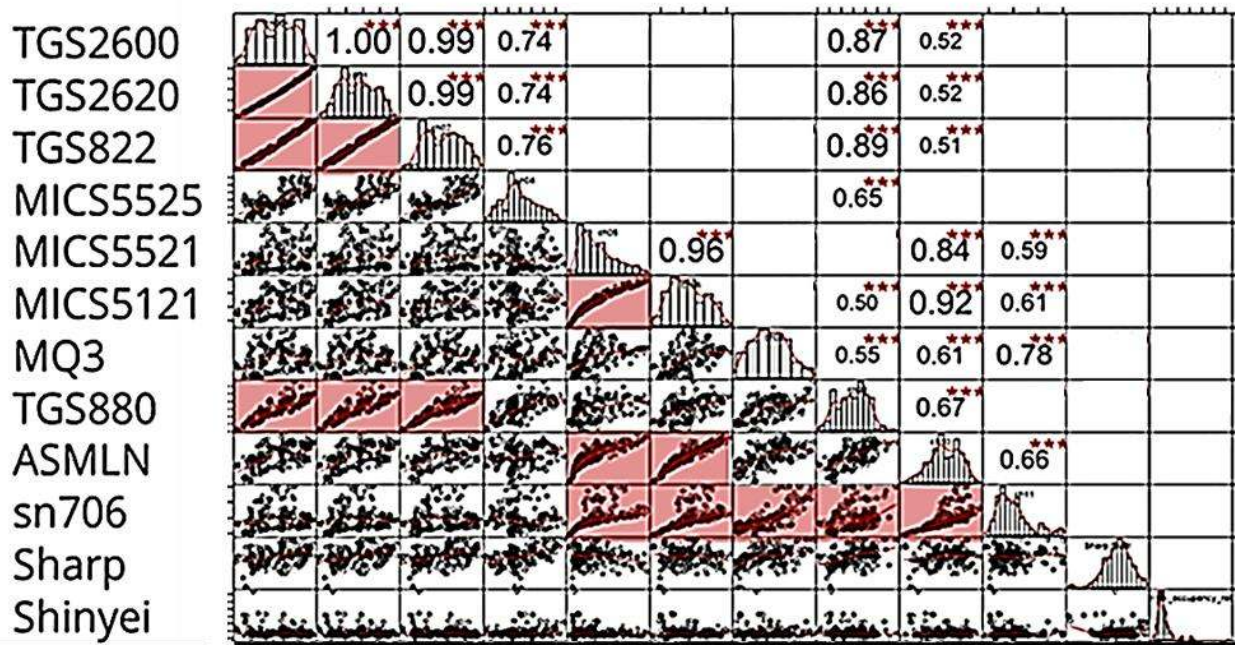


Figure 22. Assessment of strength of correlation between sensors. Pearson correlation coefficients are shown only when bigger than 0.5. The quantity of stars indicates significance as indicated by p value. The sensor data is shown in arbitrary units. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

Some sensors were found to be weakly correlated to pollutant targets and were strongly cross-responsive (i.e. TGS2600, TGS2620, TGS822). In several cases, same sensor was found to be correlated to several pollutants. For example, MICS5521 was found to be correlated to CO₂, NO_x and O₃. Other sensors were characterized by higher selectivity. For example, TGS880 showed higher correlation to CO as compared to other pollutants. Some gas sensors (i.e. TGS822) were found to be better correlated to PM_{2.5} than some actual particulate matter sensors. In addition, reference pollution monitor response distribution was found to be more skewed as compared to sensor response distribution suggesting that sensors might not be as sensitive to higher pollutant concentrations or that the sensor response was related to the log of pollutant concentration.

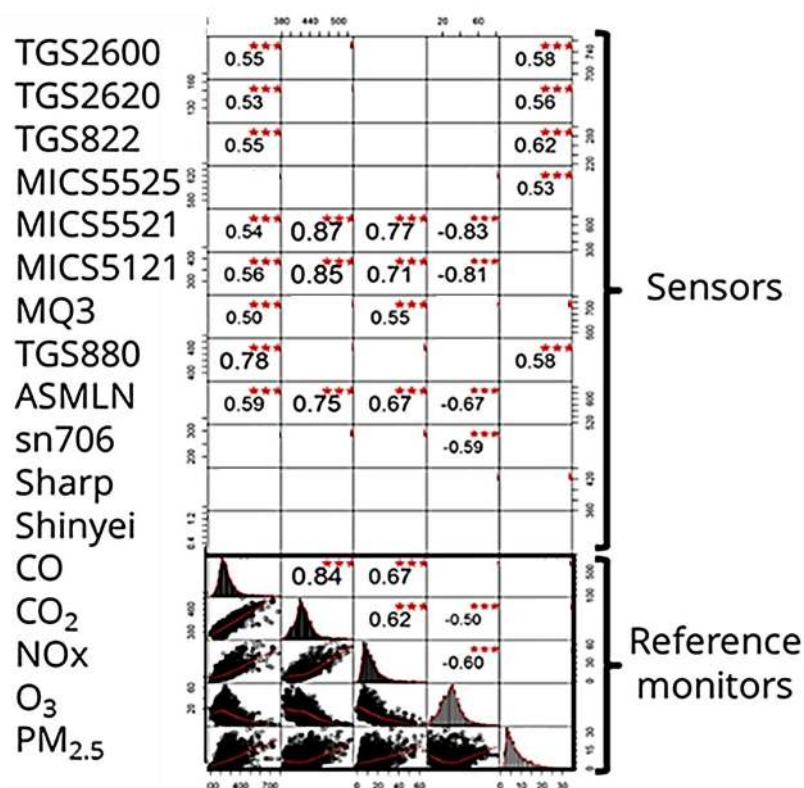


Figure 23. Assessment of strength of correlation between sensors and reference monitoring instruments for carbon monoxide, carbon dioxide, nitric oxides, ozone and particulate matter. Pearson correlation coefficients are shown. The quantity of stars indicates significance as indicated by p value. The sensor data is shown in arbitrary units, while reference data is shown in concentration units (i.e. ppb, ugm-3). Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.1.5 Relative sensor importance

Several different methods were used to evaluate sensors to select the best sensor subsets to be used in calibration models for each of the investigated air pollutants (i.e. NO_x, PM_{2.5}, CO, CO₂, O₃). Sensor subset selection is an important step for developing calibration models because the challenges and opportunities associated with sensor cross-responsiveness need to be balanced. While multivariate regression models with large number of sensor predictors can be more accurate, they can also lead to overfitting and unstable performance at locations with different pollutant compositions. Another goal

for sensor subset selection was to identify the sensors that could be excluded from AirSENCE device, thereby reducing the cost and power consumption.

The first type of analysis used to evaluate sensor predictor subsets was the analysis of standardized regression coefficients (which allows the sensor predictors to be evaluated on a common scale) and p value (which describes the degree of sensor importance significance). For the purpose of this analysis, sensors for which p values are below 0.2 (i.e. 80% confidence level) were considered significant enough to be included in the model. However, p value doesn't describe the magnitude of the effect of sensor variable on the predictive ability of the model. Therefore, standardized regression coefficients were also evaluated to check whether some of the sensors should be excluded due to the small effect they are having on model performance. Standardized regression coefficients were determined by subtracting the mean and dividing by the standard deviation of the sensor variable so that the mean is zero and the variance is 1. The standardized regression coefficients of the resulting model show by how many standard deviations a sensor reading changes per standard deviation change of pollutant concentration as measured by the reference instrument.

To further investigate the sensor subsets that result in improved model performance, a search was performed for best subsets of sensors. Four different regression search algorithms were compared: exhaustive search, forward stepwise, backward stepwise and sequential replacement. The relaimpo R package was used to compute variable importance. The subset models were evaluated using two model performance metrics – adjusted r-squared (r-squared adjusted for the number of explanatory terms in the model) and Schwartz's information criterion BIC (criterion for model selection that incorporates a penalty term for the number of parameters in the model to discourage overfitting). Both adjusted r-squared and BIC gave almost identical performance. The model subset analysis confirmed some of the sensor importance hypotheses, which were generated by standardized regression coefficient and variable significance analyses. However, sensor subset selection analysis also suggested importance of some other sensors, which were not considered important before.

One of the challenges associated with sensor variable ranking using searches based on regression model comparison is that there are many possible subsets that appear to have similar performance, which commonly results in local minimum sensor subset that does not perform best across all of the environmental conditions the AirSENCE devices might be exposed to (Gualdrón et al. 2006). Therefore, sensor variable ranking using a random forests regression algorithm was also evaluated to analyze the

relative importance of different sensors and identify candidates for exclusion. The random forest model is an ensemble learning model, which works by constructing multiple decision trees and evaluating mode of the output of those trees. This method can be used for ranking the importance of sensors by calculating mean square error (%MSE) of the prediction when the sensor is removed from the model. Since different sensor array devices result in slightly different ranking, the ranking scores were averaged over 5 devices to reduce any bias. Random forest-based variable ranking helped identify sensor candidates for exclusion.

The summary of the 3 analysis methods was used to identify the final predictor subsets for NO_x, O₃, CO, CO₂, PM_{2.5} and AQHI (Table 11). An overview of the final sensor predictor subset as well as corresponding calibration model coefficients is provided in Appendix 3: Finalized calibration model predictor subsets and coefficients. Combination of NO₂ and NO (i.e. NO_x) was used as target rather than the component gases because models with adequate accuracy for NO₂ and NO could not be developed using the output from the available sensors. The selectivity of MOS sensors is insufficient to differentiate NO₂ and NO gases. While other sensor types (i.e. electrochemical sensors) have been shown to differentiate NO_x species, they have not been investigated in this work and are beyond the scope of this thesis. Sensors that were found to be important as indicated by all 3 analysis methods were used in the final predictor subsets. Many of the sensors served as predictors for several pollutants, as expected from the cross-responsive nature of MOS sensors. For instance, temperature, relative humidity and P/N706 sensors were found to be important for most pollutant targets. Notably, the best predictors for each of the target pollutants are often not the ones expected based on the information from the sensor datasheets (Table 10). The data included in sensor manufacturer datasheets is often limited to lab studies with a small number of target pollutants. As a result, much of the sensor response properties are underexplored, which highlights the importance of comprehensive evaluation of new sensors in response to air pollution variation in ambient air. Variable selection process is summarized in Figure 24.

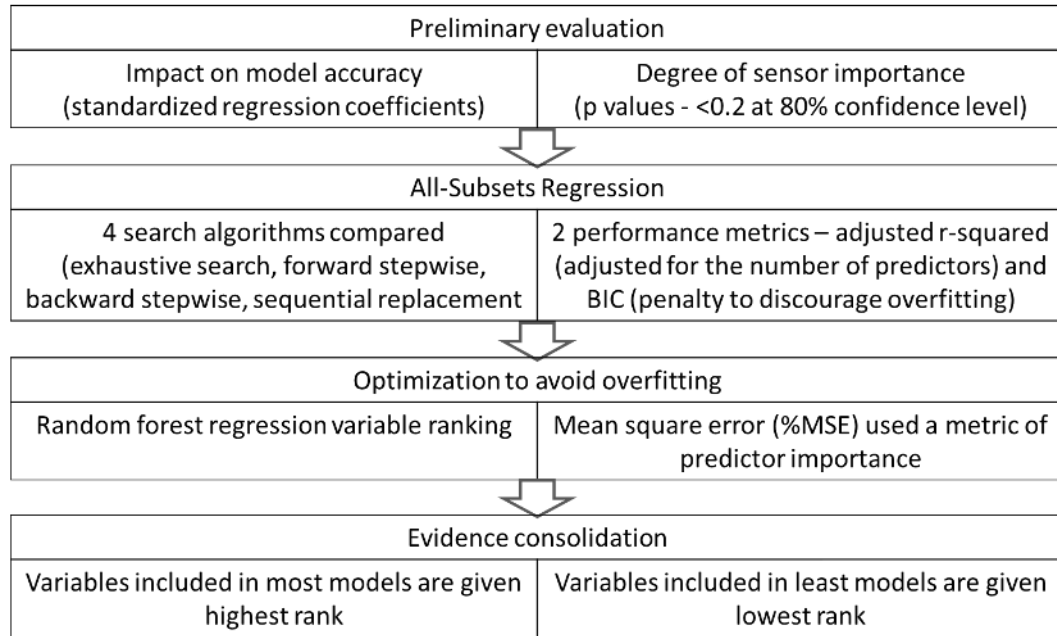


Figure 24. Variable selection process.

Table 11. Best-performing sensor vs pollutant matrix after evaluating variable importance. Pollutants of interest are listed along top row. Selected sensors are shown in green. Sensors that were found to be important as indicated by all 3 analysis methods for each pollutant were used as final predictor subsets.

Sensor Model	NOx	O ₃	CO	CO ₂	PM _{2.5}	AQHI	Selected
DHT22	X	X	X	X		X	5
TGS2600					X	X	2
TGS2620			X		X	X	3
TGS822	X		X		X	X	4
TGS880					X	X	2
P/N706	X	X	X	X	X	X	6
GP2Y1010AU0F					X	X	2
PPD42					X	X	2
Selected	3	2	4	2	7	8	

4.1.6 Site variation

As discussed earlier, sensor reproducibility is key for accurate assessment of spatial variation in pollutant levels. The analysis of inter-device sensor variation helped identify variability in sensor response to pollutants at a single site. However, deployment of devices at other location leads to new challenges. The pollution sources and therefore pollutant mixtures, chemistry and interferences often change depending on location. Sensors are often highly sensitive to these changes due to their cross-responsiveness.

The hypothesis that change in levels and types of interferences can influence change in sensor response was tested by monitoring the change in correlation strength between reference instruments and sensors at SOCAAR's downtown Toronto site, Toronto Island site and Highway 401 site (Table 12). Toronto Island site is an example of background site with low level of traffic pollution, while Highway 401 site is located close to highway and is an example of site with high levels of traffic pollution. The downtown site is close to an arterial road and is affected by medium levels of traffic pollution.

Table 12. Average, minimum and maximum pollution levels at the three sites investigated during AirSENCE deployment campaigns (April – December 2015).

Site	NOx (ppb)	O ₃ (ppb)	CO (ppb)	CO ₂ (ppm)	PM _{2.5} (ugm-3)
SOCAAR's downtown Toronto	15 (0-102)	25 (2-89)	272 (35-1039)	430 (378-593)	8 (0-148)
Toronto Island (Halans Point)	6 (0-42)	15 (0-49)	190 (59-447)	414 (398-469)	9 (1-35)
Highway 401	48 (2-266)	10 (0-35)	337 (7-1156)	389 (350-570)	11 (0-99)

It was found that the importance of predictors in estimating NOx and CO concentration changes depending on deployment location (Figure 25). Sensor cross-sensitivity results in change in correlation between sensors and pollutants. For example, the sensors TGS880, ASMLN and P/N706 were highly correlated to the NOx reference instrument at the SOCAAR site. However, the correlation strength for these sensors was significantly smaller at both the Island and Highway sites. This suggests that while these sensors are responding to the same pollutant(s), these pollutants do not necessarily include NOx. The sensors might be responding to proxy contaminants emitted by the same sources as NOx or the contaminants that are associated with NOx through chemical reactions. This result highlights the importance of calibrating AirSENCE devices at several different types of sites to ensure robust performance. The correlation between MQ3 and NOx decreased but was still the highest compared to

other sensors after the deployment, which suggests that this sensor should be used for NO_x model correction. For CO, the ASMLN and MQ3 sensors retained the highest correlation to reference monitors after deployment at the Island and Highway sites.

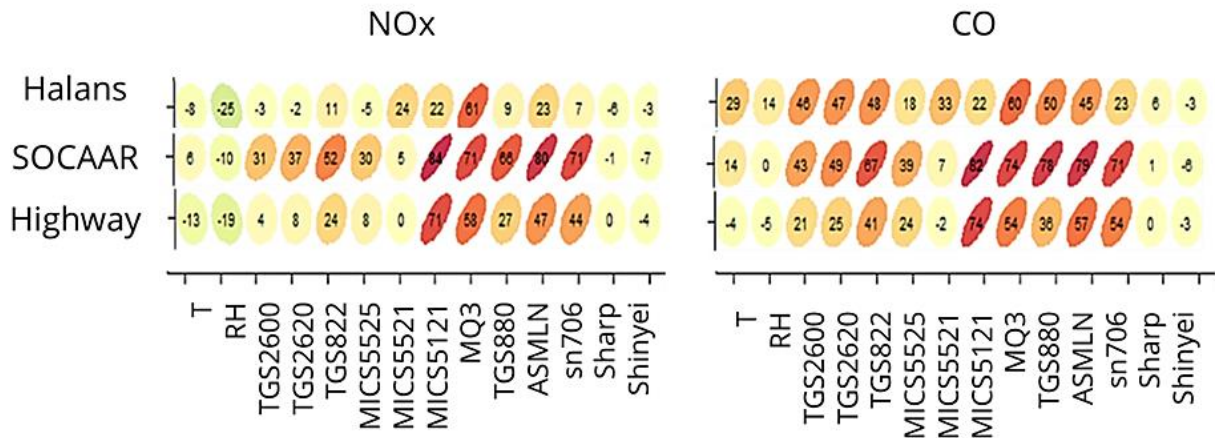


Figure 25. Pearson correlation coefficients between reference instruments and AirSENCE devices for NO_x and CO at three contrasting sites. The numbers shown are Pearson correlation coefficients multiplied by 100. The colour of ellipses and the shape indicates the degree of correlation. The negative sign indicates negative correlation. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.1.7 Sensor response drift

Drift in sensor response is another major known factor impacting sensor performance. It is important to quantify the magnitude and direction of the drift in sensor baseline and reduction in sensitivity as well as the useful operational lifetime of the sensors.

Sensor readings were first rescaled 0-1 to correct for inter-device sensor systematic offset. Sensor baseline was calculated by taking a 7-day average of rescaled sensor response. Sensor sensitivity (i.e. range) was calculated by taking the difference between maximum and minimum value for each day and averaging daily values over the course of 7 days. Averaging allowed for assessment of long-term drift while correcting for known temporal patterns (i.e. variation over the course of a day or a week). Drift in sensor response was quantified by assessing the change in the value of sensor baseline and range over time (Figure 26).

To assess the impact of drift on sensor response under more targeted conditions, the validation dataset was divided into subsets, selecting only the conditions that would have limited impact on sensor response. For example, NO_x and O₃ are known to vary seasonally and have an impact on sensor response. By selecting the data associated with lower quartiles of NO_x (0-9ppb) and O₃ (5-17ppb) over a 6-month period (2 seasons), we can extract variation in sensor response over time without confounding the analysis with any seasonality of these pollutants. Temperature (high and low extremes) and humidity (high extreme) are also known to have a large impact on sensor response. Therefore, 2nd temperature quartile and 1st humidity quartile subsets were selected for analysis. The data was subset for each condition separately since it was impossible to correct for all conditions simultaneously because these conditions didn't co-occur in the available ambient data. In addition, analysis was conducted for 3 sensor replicates separately, since sensor average is not necessarily representative. The overall impact of temporal drift was categorized as low or high by comparing the average of sensor metrics (i.e. baseline and response range) for each consecutive month and verifying the significance using a t-test.

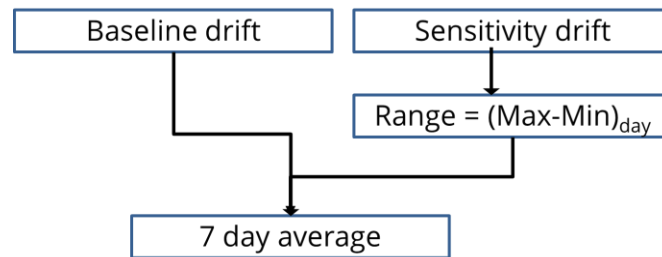


Figure 26. The process for evaluating sensor drift.

Drift impacts sensor response after 4 months of operation and sensors were significantly affected by drift in baseline and reduction in range over time (Figure 27). The P/N706 and PPD42 sensors were most effected by drift in baseline and reduction in sensitivity. Ironically, sensor drift also contributes to reduced inter-device sensor variation over time as unresponsive sensors agree well with each other.

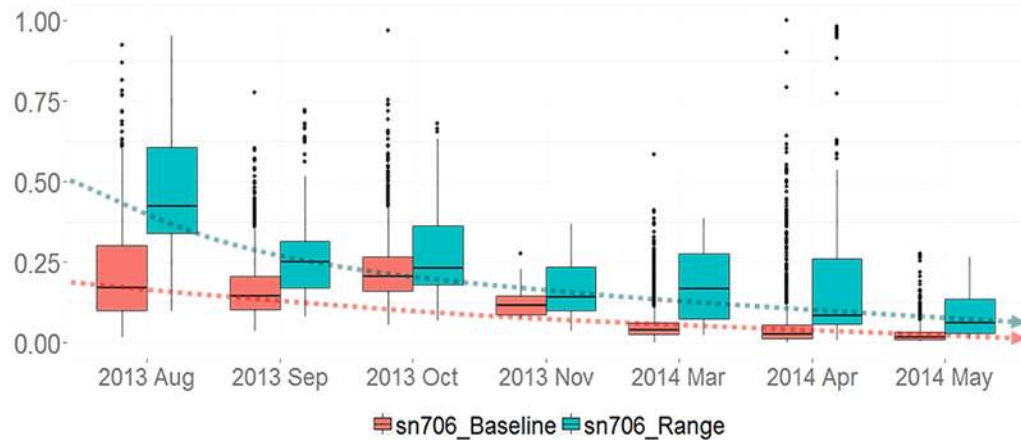


Figure 27. Change in baseline (drift) and range (sensitivity) in P/N706 sensor response (average of four devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

In general, a reduction in sensitivity was more common than drift in baseline (Table 13). In addition, all sensors that are affected by baseline drift are also affected by a reduction in sensor response range. Five sensors were not affected by drift in baseline and reduction in range. These sensors could be used for correcting the response of the sensors that are more affected by drift. For instance, earlier analysis indicated that sensors TGS822, TGS2620, TGS2600 and TGS880 are inter-correlated. Since TGS822 and TGS2620 are less affected by drift, they can be used to correct TGS2600 and TGS880 sensors for drift in response range. In addition, P/N706 is correlated with MICS5521, MICS5121 and MQ3. Therefore, sensors with larger drift can potentially be corrected using output of sensors with less drift.

For example, sensor MQ3 shows less baseline drift compared to sensor P/N706 (Figure 28). However, the two sensors are correlated, allowing for baseline drift correction of the P/N706 sensor response (Figure 29). The baseline drift correction was accomplished by subtracting P/N706 sensor baseline and adding MQ3 sensor baseline corrected for the difference between P/N706 and MQ3 sensors.

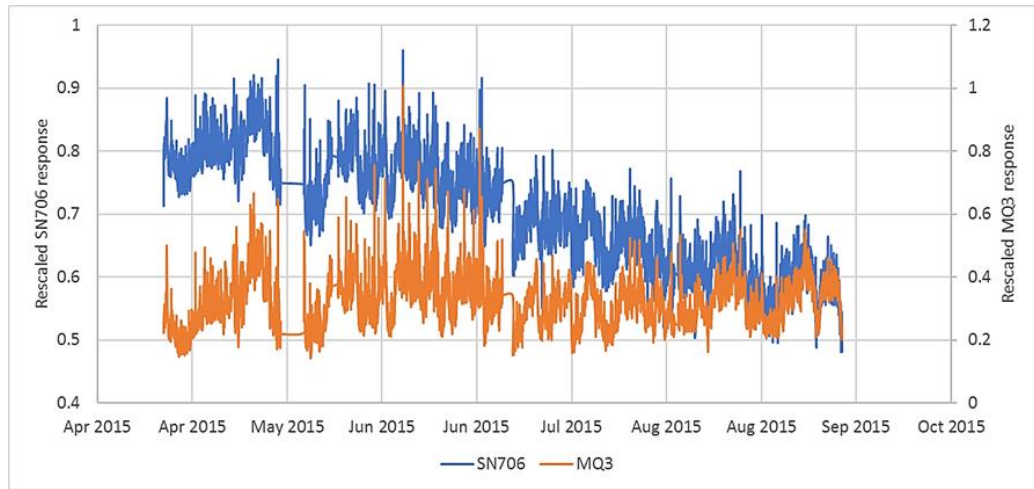


Figure 28. Comparison of drift for sensor with high sensor drift (P/N706) and low sensor drift (MQ3). Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.

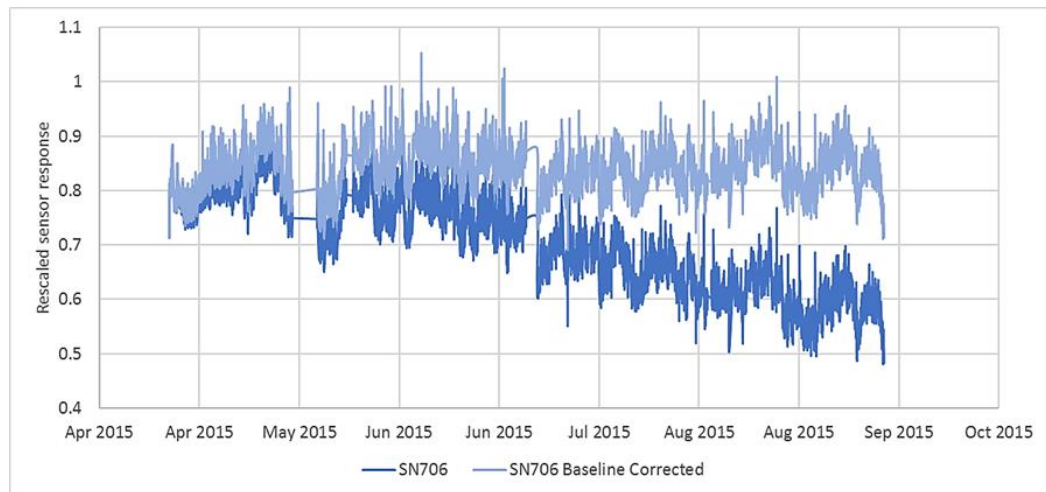


Figure 29. Sensor baseline drift correction for P/N706 using MQ3 sensor data. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.

Table 13. The impact of baseline and range (sensitivity) drift over time on the response of all the tested sensors.

	Baseline impact
--	-----------------

Sensitivity (range) impact		Low	High
	Low	TGS822, TGS 2620, MICS5121, TGS825, MQ3	
	High	GP2Y1010AU0F, TGS2600, MICS5525, MICS5521, MQ135, TGS880	P/N706, PPD42, MICS2610

4.1.8 Summary

In summary, sensors were found to be affected by high inter-device sensor variability. A major systematic error contributing to inter-device sensor variation is the difference in the degree of sensor response relative to the reference (i.e. sensitivity) for different sensor replicates. High temperature and high CO and O₃ concentrations also contribute to variation between sensor replicates and should be accounted for.

When sensor importance was evaluated, it was found that same types of sensors are often correlated with multiple pollutants and predictor importance analysis suggests that same sensors should be used in calibration models for several pollutants. When sensor signal distribution is compared to that of reference instruments, sensor distributions were found to be less left-skewed suggesting reduced sensitivity to higher pollutant levels. Sensor response non-linearity was found to be a major cause and sensor response transformation was identified as a potential solution. Calibration of sensors using higher pollution level is another potential solution as it would allow for development of a separate calibration model for higher pollutant levels rather than relying on extrapolation. This solution is explored in Section 4.4 of this document.

It was also found that many sensors are cross-responsive. Cross-sensitivity results in change in correlation between sensors and pollutants at different sites and suggests the importance of calibrating the AirSENCE devices at multiple sites. Finally, drift was found to have a large impact on sensor performance after 4 months of operation. Drift in sensitivity affects more sensors than drift in baseline.

4.2 Pollutant model performance

The second major aim of this research work was to evaluate the performance of calibration models that relate sensor measurements to pollutant concentration. The performance of calibration models for NO_x,

O₃, CO, CO₂, PM_{2.5} combining different types of sensors was evaluated and devices were characterized with respect to accuracy, precision, sensitivity and limit of detection. Devices, in the context of this work, are defined as arrays of 15 non-identical sensors (Table 10) connected to the same microcontroller and power supply. Each device represents a complete sensor array that has been independently calibrated against the references monitors. The resulting calibration models for each target pollutant are the same with respect to sensors used but vary with respect to the calibration model coefficients. The variation of performance with changes in pollutant concentration was also explored as well as the factors contributing to lack of accuracy. Performance of AQHI models was also evaluated. The following research questions were of interest:

Calibration model performance

1. What is the impact of combining different types of sensors on calibration model performance?
2. Which pollutants can be predicted with highest (and lowest) accuracy and precision?
3. What is the sensitivity of the sensor array towards each pollutant and how does this inform the applications for which the sensor array can be used?
4. How linear is the sensor array response towards each pollutant?
5. What is the lowest concentration at which each pollutant can be detected and quantified?

Factors affecting calibration model performance

6. How does the accuracy and precision of the prediction vary with pollutant concentration?
7. What are the major factors affecting the calibration model accuracy?
8. How does variation in reference residuals compare to that of inter-device residuals?

Variability in performance across devices and sites

9. How does inter-device variation in pollutant estimates compare with their accuracy?
10. How does an Air Quality Health Index (AQHI) estimates derived directly using model fitting compare to those calculated from pollutant estimates? How does performance vary across sites?

4.2.1 *Multivariable models*

Due to cross-responsive nature of sensors, combining selective as well as supporting sensor characteristics with environmental variables was postulated to lead to more robust models with higher sensitivity and reduced impact of interferences. This premise was tested by combining the sensors selected by variable importance subset search algorithms to derive multiple linear regression (MLR)

models for NO_x, O₃, CO, CO₂ and PM_{2.5}. The R package “Stats” was used to build and evaluate the models and MLR regression coefficients for each sensor device in response to each pollutant were determined. Although sensor data was baseline-corrected, there was still large variability in sensitivity of sensor replicates. As a result, the models for each device were formulated using the same predictors (sensors) but with device specific regression coefficients associated with each of the predictors. To evaluate performance of the MLR models, measurements were analyzed at 15min time resolution to capture full-scale diurnal variation of pollutant levels, while averaging out any effects irrelevant for this analysis (ex. differences in time scale of pollutant dispersion and other short time scale effects). Analysis of residual time series was used to for initial evaluation of device model. Residuals between each pollutant model observation and multiple device pollutant model average was used to assess how comparable the devices are to each other and the residual between each pollutant model observation and reference monitor was used for initial assessment of accuracy.

AirSENCE devices Rev1 were found to track the reference instrument well but underestimated some pollutant peak levels (Figure 30). The analysis also showed that combining sensor response variability (dimensionality) as well as temperature and humidity measurements lead to improved accuracy of pollutant models. Due to cross-response nature of some sensors (i.e. O₃ metal oxide semiconductor sensors), addition of other sensor predictors (i.e. NO_x sensors) was necessary for correcting the sensor response and improving the accuracy of pollutant estimates.

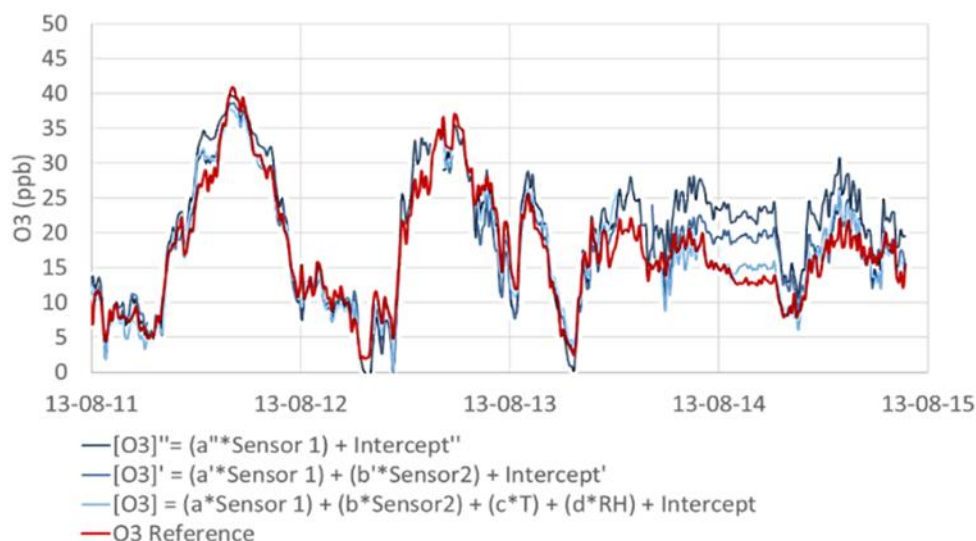


Figure 30. The performance of ozone MLR model improves as predictors are added to the model. The output of the ozone sensor model [O₃]'' is compared to a model making use of both ozone and NO_x

sensor predictors [O_3]’ as well as a model combining ozone, NO_x , temperature and humidity sensors [O_3]. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

4.2.2 *Model performance*

Evaluation of quantitative performance metrics was important for gaining a full understanding of limitations of pollution monitoring using AirSENCE devices. Accuracy and precision helped assess the sensor device performance characteristics and indicated which models need to be optimized. The evaluation of linearity and sensitivity helped verify whether linear regression models explain all variability in sensor response. Analysis of limit of detection and limit of quantification as well as inter-device variation helped identify the types of applications the AirSENCE devices could be suitable for.

Regression model performance was evaluated using a set of regression metrics (i.e. accuracy, precision, linearity, sensitivity, limit of detection and limit of quantitation). All the definitions for sensor performance metrics were taken from the “17025” standard developed by International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC) as well as the American Society for Testing and Materials (ASTM) documentation (ISO 2005; ASTM International 2016; IEC 2017). All the metrics were evaluated for a set of 4 devices to ensure variability between devices is captured. A MLR model was built for each device for each of the five pollutant targets: NO_x , O_3 , CO, CO_2 , $PM_{2.5}$. The regression metrics were evaluated on the training set to determine the best-case performance of the MLR models.

Precision was defined as the agreement (degree of scatter) between a series of measurements of a sample. Precision can change depending on pollutant concentration, sensor operating conditions and environmental variables (i.e. temperature, humidity). Different types of analyses and metrics were used to assess the precision of the sensors:

- repeatability (evaluated under the same pollution source profile and environmental conditions),
- intermediate precision (evaluated under same source profile but different environmental conditions)
- reproducibility (evaluated for different source profiles or mixtures of pollutants).

While sensor repeatability was evaluated during calibration (Section 4.2), intermediate precision was evaluated during long-term device testing (Section 4.3.1) at SOCAAR and reproducibility was evaluated when the devices were tested at other sites (Section 4.3.4). Analysis of reproducibility informs sensor response robustness and ruggedness. For the purpose of calibration model performance evaluation, sensor precision was expressed as relative standard deviation (%RSD) (Equation 1). Relative standard deviation was calculated for each device by dividing each device model output by average sensor model output at that concentration. Precision %RSD represents variation of one device in its ability to reproducibly measure a specific pollutant concentration

Equation 1. Equation for precision or inter-device variation. %RSD is % relative standard deviation and s is the standard deviation of the sensor device replicate measurements. For precision measurement, \bar{x} is the average pollutant estimate for a concentration level for one device, while for inter-device variation, \bar{x} is average pollutant estimate for sensor device replicates.

$$\%RSD = \frac{s}{\bar{x}} \times 100$$

Precision also informs LOD – the smallest change in concentration that can be detected. Limit of detection (LOD) was determined as the lowest amount of pollutant, which could be detected. It was quantified by identifying the point for which measurement value was larger than the uncertainty associated with it. Limit of detection was calculated by multiplying standard deviation of baseline response by 3. Standard deviation of baseline was calculated using standard deviation of lower end of calibration range (i.e. zero air).

Limit of quantitation (LOQ) is defined as the lowest amount of pollutant which can be quantified with acceptable accuracy and precision. Limit of quantitation was calculated using two methods: 1) by multiplying standard deviation of baseline response (calculated the same way as LOD) by 10, and 2) by finding the lowest concentration, at which the measured value does not exceed 30% RSD (precision) and 30% absolute % error (accuracy) for the actual measured pollutant concentration range. The biggest value out of the two calculated values was taken as limit of quantitation. Two methods were used to calculate LOQ because even though low standard deviation may indicate very low LOQ, the actual estimation can be higher depending on the measured pollutant concentration range.

A related metric to precision is inter-device variation, which is defined as variation between device replicates for a specific pollutant concentration. For the purpose of calibration model performance evaluation, inter-device variation was expressed as relative standard deviation (%RSD). Inter-device %RSD was calculated by dividing each device model output by average sensor model output of 5 devices for each observation.

Accuracy was defined as the difference between pollutant concentration estimated by sensor device and accepted reference value (Equation 2). Sensor device accuracy varied depending on pollutant concentration and environmental conditions. In addition, accuracy changed over time as a result of baseline drift and sensitivity decline due to sensor aging. Accuracy was quantified as mean absolute percentage error between sensor measurement and reference measurement. It was calculated by finding the difference between reference monitor output and sensor device model output and dividing it by reference monitor output. The value was then converted to an absolute percentage. To aid with comparisons between different types of pollutants, the reference residual was defined as average residual expressed as % full scale (i.e. residual divided by the difference between maximum and minimum pollutant level).

Equation 2. Accuracy equation. A is average absolute % Error, A_t is the target pollutant reference value, F_t is sensor measurement and n is the number of observations.

$$A = \frac{100\%}{n} \sum_{t=1}^n \frac{A_t - F_t}{A_t}$$

Sensitivity is defined as change in sensor device regression model output per unit change in pollutant concentration (Equation 3). It was calculated by finding the slope of the relationship between sensor device regression output and reference monitor output.

Equation 3. Sensitivity equation. M is sensitivity, Δy is change sensor device regression model output, Δx is change in pollutant concentration during the same period.

$$m = \frac{\Delta y}{\Delta x}$$

Linearity is defined as degree to which unique sensor observations deviate from the linear regression line (Equation 4). Linearity was defined as good for the pollutant range where regression line residual (deviation from regression line) did not significantly exceed sensor precision. Linearity was expressed quantitatively as % best fit nonlinearity, which is the ratio of the maximum output deviation from the best-fit line divided by the full-scale output, specified as a percentage. Since ambient air pollution exposure was used to calibrate the AirSENCE devices, the pollutant concentration range for which linearity could be assessed was limited by the ambient pollution concentration range. In addition, linearity was qualitatively evaluated using 2 plots: 1) sensor vs reference scatter plot, 2) sensor residual vs reference.

Equation 4. Nonlinearity equation. %NL is percent best fit nonlinearity, Dmax is maximum output deviation from the best-fit line and FS is full-scale output.

$$\% \text{ NL} = D_{\text{max}} / F_S * 100$$

MLR models for CO and CO₂ were found to demonstrate the best performance in terms of precision and absolute accuracy (Table 14). However, it is important to note that when the average residual was normalized by pollutant full scale variation, the accuracy of MLR models for these pollutants were comparable to NO_x, O₃ and PM_{2.5}. The higher average levels relative to the ranges of CO and CO₂ is a major reason for this discrepancy.

Some of the contributors to lower accuracy in NO_x, O₃ and PM_{2.5} sensor device measurements are sensor response nonlinearity and lower sensitivity in MLR models for these pollutants (Table 14). In addition, the limit of quantitation is also higher for NO_x, O₃ and PM_{2.5}, suggesting that one of the causes for lower accuracy is lower signal-to-noise ratio.

Table 14. Calibration performance metrics for inorganic gases and particulate matter. Precision and accuracy were evaluated using pollutant concentration range excluding the data below LOQ. Average value for each of the performance metrics is shown for the set of 4 devices as well as standard deviation, which represents variation between devices. (* LOQ values were evaluated using two methods described in the text). Original data was collected between August 15, 2013 and May 10, 2014,

sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

	NO _x	O ₃	CO	CO ₂	PM _{2.5}
Pollutant range during calibration week	1.7 - 62 ppb	1.9 - 75 ppb	201- 729 ppb	371- 453 ppm	0.5 – 33 ugm-3
Mean ± SD	27 ± 7	34 ± 9	340 ± 78	395 ± 15	17 ± 4
Quartiles (Q1, Q2, Q3, Q4)	7, 11, 15, 62 ppb	15, 21, 31, 75 ppb	287, 333, 38, 729 ppb	385, 393, 403, 453 ppm	4, 6, 9, 33 ugm-3
Precision - Repeatability (%RSD)	22 ± 10%	14 ± 10%	9 ± 4%	0.7 ± 0.6%	13 ± 8%
Accuracy (abs % error, residual as % full scale)	29 ± 16% 23%	26 ± 9% 25%	7 ± 10% 29%	2 ± 1% 35%	38 ± 14% 24%
% Nonlinearity	55 ± 7%	63 ± 17%	32 ± 5%	8 ± 2%	55 ± 4%
Sensitivity (Δ sensor/ Δ reference)	0.83 ± 0.03	0.9 ± 0.03	0.97 ± 0.008	0.99 ± 0.0003	0.85 ± 0.004
LOD	5 ± 1 ppb	11 ± 1 ppb	65 ± 20 ppb	11 ± 1 ppm	3.5 ± 0.5 ugm-3
LOQ * (method 1, method 2)	18 ± 5 ppb 18 ppb	37 ± 3 ppb 20 ppb	217 ± 68 ppb 250 ppb	35 ± 4 ppm 371 ppm	12 ± 2 ugm-3 15 ugm-3

4.2.3 Accuracy and precision

Assessment of the variability of accuracy and precision with change in pollutant concentration is important for understanding the usable range of AirSENCE devices as well as the characteristics of their

performance. Magnitude and direction of change in accuracy and precision in response to pollutant concentration also informs the performance of AirSENCE devices over time and space.

It was found that precision increases as the pollutant concentration increases, while accuracy is highest for mid-range ambient pollutant concentrations, as indicated with the sample ozone regression model data (Figure 31). Data for other pollutants can be found in Appendix 4: Precision and accuracy analysis. The low level of precision at low pollutant concentration is caused by high variability in sensor measurements at pollutant levels close to the limit of detection. The decrease in accuracy at higher pollutant concentration is associated with negative % error or underestimation of pollutant concentration by AirSENCE devices. The underestimation of pollutant levels is partially caused by skewed pollutant concentration distributions, resulting in the majority of the training data corresponding to low-medium pollutant levels. Another factor contributing to underestimation of pollutant levels is nonlinearity in sensor measurements. Nonlinearity leads to reduced performance of MLR model when the model is used for extrapolating the pollutant concentrations associated with sensor response outside of the training range. Methods for compensating for these issues are explored in section 4.4.

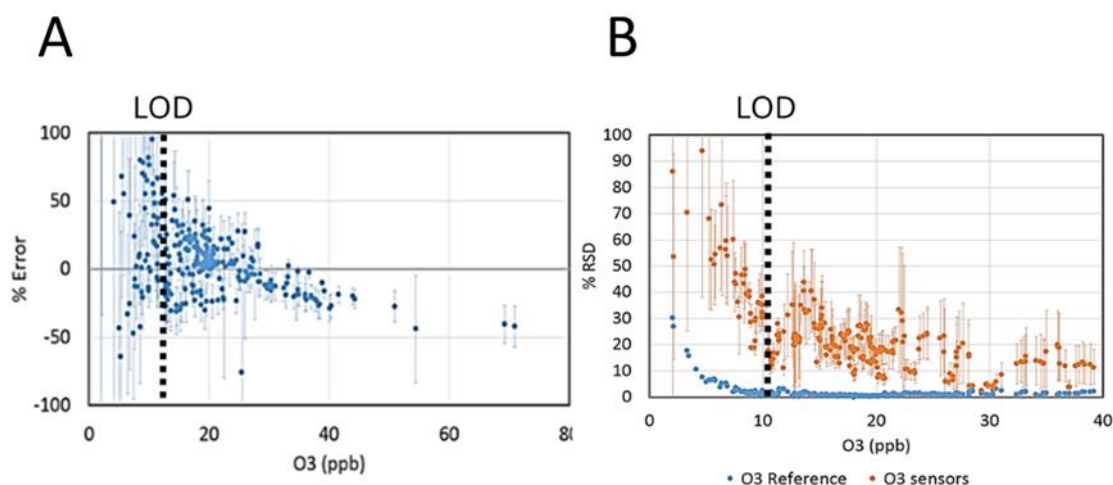


Figure 31. Evaluation of accuracy and precision of ozone model with change in ozone concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of ozone concentration over the ozone concentration range during the calibration week. Error bars represent variation between devices in predicting ozone concentration (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices ($n=4$) as a function of ozone

concentration. Vertical lines represent limit of detection (LOD). Error bars represent variation between devices in predicting ozone concentration (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

4.2.4 *Model nonlinearity affects performance*

One of the factors contributing to the high % error for $\text{PM}_{2.5}$, NO_x and O_3 was the nonlinearity of sensor response at higher pollutant concentrations. Pollutant slope tends to decrease with increased pollutant concentration and MLR performance decreases. To further understand the impact of nonlinearity on accuracy of AirSENCE devices, the impact of nonlinearity was compared for different target pollutants. The impact of sensor response nonlinearity on resolving temporal variation in pollutant concentration was also investigated.

Nonlinearity was found to cause underestimation of peak pollutant levels because of failure to capture some of the short-term changes in pollutant levels. However, the AirSENCE devices were still able to capture long-term temporal variation well (Figure 32).

While deviation from line of best fit rarely exceeded 1 standard deviation, the residuals were still significant for higher pollutant concentrations. Sensor response nonlinearity and underestimation of higher pollutant concentration is a recurring challenge for low cost sensors (Zampolli et al. 2004; Srivastava 2003; Peterson et al. 2017). There are several factors that might contribute to underestimation of pollutant levels. The reactive MOS surface can only accommodate only a certain number of molecular interactions at a given time. Therefore, depending on kinetics of the pollutant reaction on sensor surface, sensor device response towards higher pollutant levels could be subject to saturation (Helwig et al. 2009). Inaccuracy at high pollutant levels can also be caused by the need for regression model extrapolation when the measured pollutant concentration range exceeds that of the calibration dataset. Another cause for nonlinearity is confounding effects from other environmental variables and interferences which might co-occur with pollutant peak levels (Tian et al. 2016; Smith et al. 2017). It has been shown that interference from environmental factors like ambient air humidity can be large, leading to a change in sensor response by as much as a factor of five. The effects of the interferences were also found to be inconsistent across sensor replicates, varying by up to a factor of two for any given humidity value.

When regression model is optimized for a specific set of observations, overfitting may occur resulting in non-generalizable model output for data outside of the parameter range of the training data. Based on the earlier assessment of relative importance of the model predictors (4.1.5), the P/N706 sensor (a major predictor for NO_x and O₃ models) and GP2Y1010AU0F sensor (a major predictor for PM_{2.5} models) are responsible for the nonlinearity in model estimates. Nonlinear regression models (He et al. 2017) and variable transformation may help improve calibration model performance at higher pollutant levels. These topics will be explored in Section 4.4.

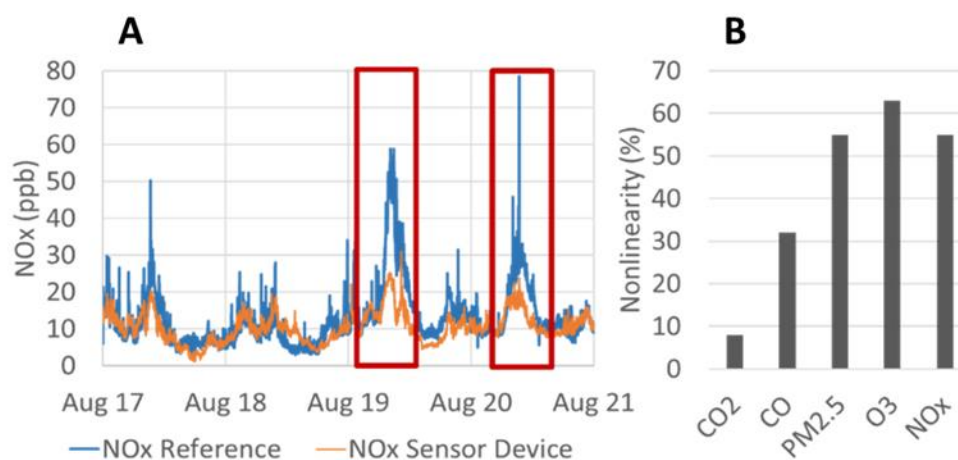


Figure 32. Model nonlinearity and its effect on pollutant models. Time series plot of NO_x MLR model output indicates nonlinearity at higher NO_x levels (A) and % nonlinearity calculation suggests that NO_x and O₃ are affected by nonlinearity the most (B). Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

4.2.5 Inter-device performance variation

Combining different types of sensors improved the regression model performance. However, one of the challenges associated with multi-sensor regression models was the requirement for calibration of each individual device due to high variability in response of sensor replicates. In fact, the MLR predictors that had greatest influence on model performance were also associated with highest inter-device sensor variation. Therefore, device-specific calibration is essential for reliable analysis. While device-specific calibration corrects for some of the variation between devices, its quantification and comparison to the

reference device residual is key for to enable spatial variability analysis. Low inter-device variability increases the confidence that the variation between devices in sensor network is real (e.g. due to actual spatial variation) rather than variability between devices.

To investigate the extent of the variation in sensor response between identical AirSENCE devices, the MLR model performance was evaluated for a set of device replicates for each pollutant of interest. The discrepancy between AirSENCE devices was compared to the discrepancy between AirSENCE devices and reference instrument for each pollutant to evaluate whether the scale and ranking of pollutants is comparable for these metrics (Figure 33). To aid with comparisons between different types of pollutants, the reference residual was defined as average residual expressed as % full scale.

Five AirSENCE devices were calibrated and tested by comparison to reference pollutant monitors at the downtown SOCAAR site. The variation between sensor device pollutant estimates and reference measurements was found to be bigger than the variation between AirSENCE devices (Figure 33). While the reference residual is above 20% of full scale for most devices, the inter-device mean %RSD is smaller than 15%. Relatively low variation between devices suggests that even when the AirSENCE devices might not be able to accurately estimate pollutant concentration, they can still be used for comparing sites. In addition, the reference residual is highest for CO and CO₂, while variation between devices is highest for O₃ and NO_x. Even though the average reference residual relative to pollutant full scale is high for CO, the devices show very similar response suggesting that CO predictors might be useful for correcting the NO_x model output at sites where these pollutants are generated by the same major source (i.e. traffic).

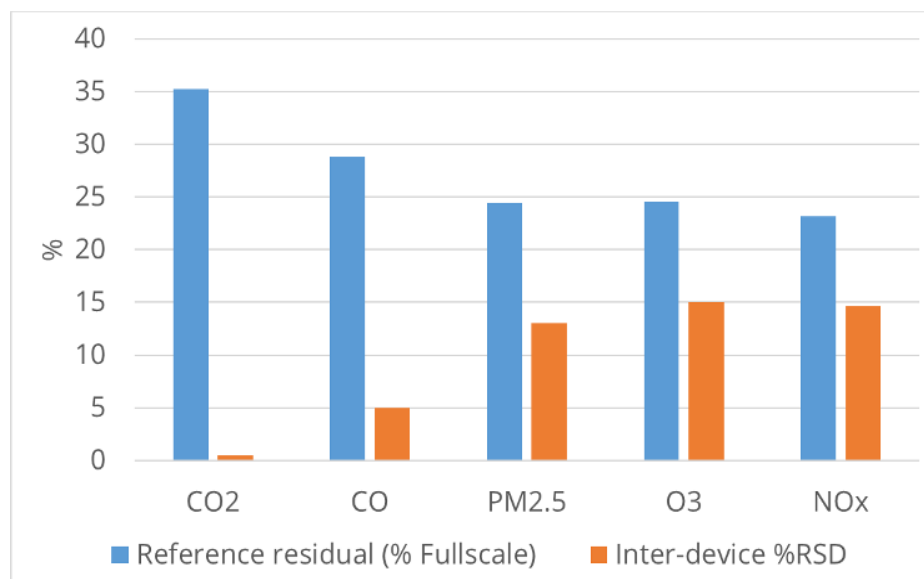


Figure 33. Comparison for inter-device variation and reference residual. Data for five devices is shown. The regression models were calibrated and tested at the same site using the same reference air pollution monitors. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

Another key finding was that while reference residual increases with pollutant concentration, the inter-device variation does not change with pollutant concentration (Figure 34). This finding is important for the analysis of sensor network data. While sensor devices at different locations might appear to have consistent response, the spatial variation of air pollution could be underestimated due to reduced regression model accuracy at high pollution levels.

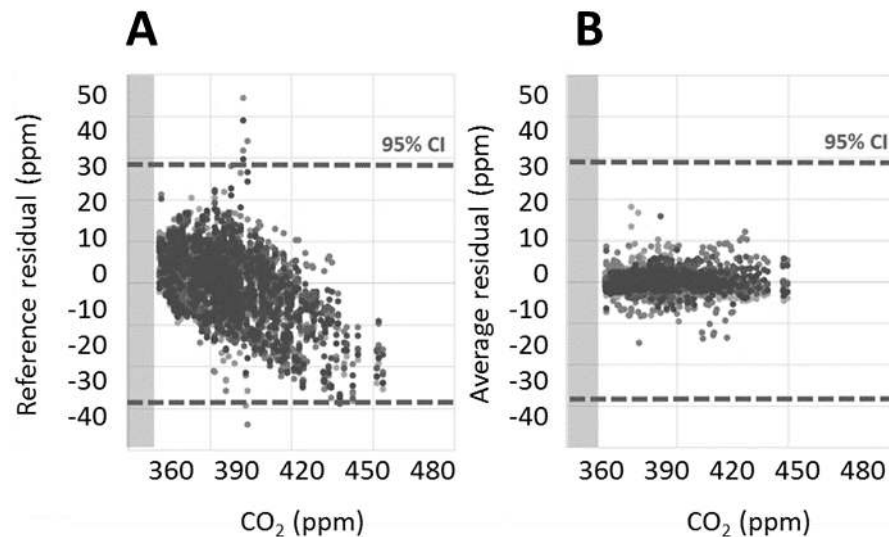


Figure 34. Analysis of residuals between sensor device MLR models and reference instrument (A) compared to residuals between sensor device MLR models and an average sensor MLR model for 5 devices (B). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

4.2.6 Multi-pollutant patterns

Pollution mixture analysis makes it possible to estimate combined metrics related to health impact. Using sensor measurements for estimation of air pollution indices enables more succinct communication of pollution data. Indices like Air Quality Health Index (AQHI) also combine the data for

multiple pollutants into a single metric. Since sensors are cross-responsive, a metric combining the measurement of several different pollutants should be easier to estimate than discrete pollutants. Air quality health index is calculated using O_3 , NO_2 and $PM_{2.5}$ pollutant data (Equation 5).

Equation 5. Air Quality Health Index calculation equation.

$$AQHI = \left(\frac{1000}{10.4} \right) \times [(e^{0.000537 \times O_3} - 1) + (e^{0.000871 \times NO_2} - 1) + (e^{0.000487 \times PM_{2.5}} - 1)]$$

There are two methods for deriving AQHI using AirSENCE devices: 1) development of MLR model for AQHI using reference AQHI calculated using reference monitor data and sensor variables as predictors, and 2) calculation of the index indirectly using the sensor device estimates for O_3 , NO_x , $PM_{2.5}$ in the AQHI formula. AQHI calculation using Equation 5 is a standard way to derive the index. The formula is simple but requires data for the NO_2 , O_3 and $PM_{2.5}$ concentrations. Three separate regression models would therefore need to be developed to derive AQHI using formula calculation method, resulting in increased analysis complexity and uncertainty due to propagation of errors from each of the models. Deriving AQHI directly from sensor predictors could reduce the analysis complexity and errors associated with the AQHI estimates but could be more impacted by sensor noise and signal artefacts. To evaluate reproducibility, the performance of AirSENCE devices in estimating AQHI was evaluated at a medium traffic site (SOCAAR), high traffic site (Highway 401) and low traffic site (Halan's point at the Toronto Island).

The analysis revealed that while both of the methods for estimating AQHI performed well at SOCAAR, the direct method (i.e. AQHI derived directly from sensor variables) performed much better when the devices were deployed at the island and highway locations (Figure 35). Model fitting method was also found to be more accurate in evaluating larger AQHI range compared to calculation method (Table 15), suggesting that AQHI model performs better at extrapolation. Sensors also are cross-responsive and nonlinear which constrains their ability to accurately resolve individual pollutant levels and contributes to some measurement errors. When AQHI is calculated using sensor device estimates for O_3 , NO_x and $PM_{2.5}$, the errors associated with estimating these individual pollutants propagate. Since the errors are highest when the AirSENCE devices are validated at a new site (due to different pollutant levels and interferences), the errors in AQHI measurements are significantly higher for calculated AQHI values compared to errors arising from a direct AQHI regression model. As a result, reproducibility is higher for

model fitting method compared to calculation method. It is also expected that model fitting method would perform better for long-term AQHI monitoring studies, but this is a topic for future work as more data is needed to test this hypothesis.

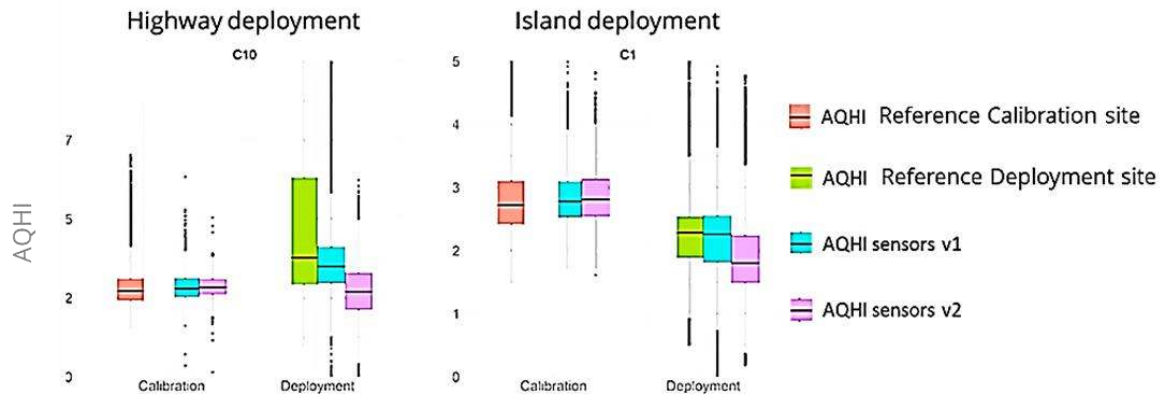


Figure 35. Comparison of performance of two Air Quality Health Index (AQHI) estimation methods at SOCAAR, Highway and Island sites. AQHI sensors v1 refers to AQHI MLR model calculated directly using the sensor data, while AQHI sensors v2 refers to AQHI measurements derived from using AirSENCE estimates for O_3 , NO_2 and $PM_{2.5}$ in the AQHI equation. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

Table 15. Standard deviation values associated two Air Quality Health Index (AQHI) estimation methods at SOCAAR, Highway and Island sites. AQHI sensors v1 refers to AQHI MLR model calculated directly using the sensor data, while AQHI sensors v2 refers to AQHI measurements derived from using AirSENCE estimates for O_3 , NO_2 and $PM_{2.5}$ in the AQHI equation. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

Location	AQHI variable	Standard deviation
SOCAAR	AQHI Reference	0.7
	AQHI Sensors v1	0.3
	AQHI Sensors v2	0.1
Island	AQHI Reference	0.7
	AQHI Sensors v1	0.5
	AQHI Sensors v2	0.2

Highway	AQHI Reference	3.4
	AQHI Sensors v1	2.3
	AQHI Sensors v2	1.9

4.2.7 Summary

In summary, MLR models using multiple sensors provided more accurate pollutant concentrations than using single sensors. A combination of different sensors is needed to determine pollutant concentration. NO_x and PM pollutant models suffer from the low precision and accuracy. While precision improves with pollutant concentration, accuracy is highest for mid-range pollutant levels. Nonlinearity is one of the major contributors to low accuracy. Evidence suggests that variable transformation or non-linear regression model development might improve performance. Difference between reference monitors and AirSENCE devices is higher than the difference between device replicates, which doesn't exceed 15%. Air quality health index can be quantified and is useful for resolving spatial and temporal pollution patterns. Its estimation by direct regression model fitting leads to a more accurate result compared to indirect calculation.

4.3 Reproducibility & the impact of interferences

Majority of studies in the literature use short-term calibration at a single site as a method for deriving calibration models and rely on these models to accurately estimate pollutant variation over space and long periods of time (Piedrahita et al. 2014; Gao, Cao, and Seto 2015). However, the assessment of device performance at the calibration site for a period of time less than a month provides only a limited snapshot of device performance (Castell et al. 2017; Spinelle, Gerboles, and Aleixandre 2013). It helps establish the best-case performance, but sensor have been shown to be cross-responsive and sensitive to drift. To be useful for air pollution monitoring applications, sensor device needs to not only show acceptable performance in measuring pollutant levels, but also good long-term reliability. Therefore, a key step in sensor device performance analysis pipeline is the assessment of consistent performance over time and in response to changes in environmental conditions and pollution source composition.

Key metrics that define device performance are:

- Change in baseline performance
- Change in range (i.e. ability to measure full span of response)

- Inter-device variation (i.e. %RSD)

Consistent baseline suggests consistency in sensor device response to low pollutant levels, while consistent range is a proxy for robust sensitivity performance. Stable inter-device variation suggests that the ability of sensor device is able to accurately assess of spatial variation of pollution stays constant.

Therefore, the impact of drift on device performance and inter-device variation was investigated to establish the long-term device performance. The impact of interferences on model performance was also analyzed and the influence of factors such as changes in temperature, humidity and pollution concentration on MLR model baseline and range was assessed. The reliability of model performance at sites with different levels of traffic was also investigated.

The following questions were addressed:

Impact of drift over time

1. What is the impact of drift on sensor device estimate of pollutant baseline, range and inter-device variation?
2. What is the change in relationship between pollutant levels estimated by AirSENCE devices over time?

Impact of interferences

3. What is the impact of interferences on sensor device response?
4. How does temperature and humidity impact the baseline, range and inter-device variation of sensor device regression models?
5. How does change in pollutant concentration influence sensor device performance and inter-device variation? What are the most influential factors with respect to the number of sensor affected?

Impact of site variability

6. How is the sensor device performance affected when devices are deployed at a site with different pollution concentration?

Assessment of the impact of drift and changes in response to environmental variables requires long-term monitoring for sufficient data to be collected. The long-term performance of AirSENCE devices was tested over a period of 9 months during deployment at SOCAAR site. In order to evaluate the long-term performance of sensors, the devices were co-located at the SOCAAR lab. Evaluation was conducted only

for AirSENCE devices that showed acceptable performance during calibration in order to ensure consistent benchmark.

The impact of several different factors on sensor performance was assessed:

- Drift over time
- Variation in environmental conditions (temperature, humidity)
- Variation in pollutant concentration (i.e. NO_x, O₃, CO, CO₂, PM_{2.5})
- Level of pollution at the site

Baseline drift was quantified by taking a 7-day average of regression model estimates and evaluating the change in weekly averages over time. Change in sensitivity was quantified by calculating the change in regression model range. Regression model range was calculated by finding the difference between minimum and maximum pollutant concentration over the course of a day and averaging the daily range values over 7-day period. Averaging over 7 days allowed us to assess the impact of drift and environmental factors on sensor device performance, while averaging out known temporal patterns (i.e. variation of pollutant levels over the course of a day or a week).

The goal of the analysis was to quantify change in sensor performance. However, the reference instrument is also expected to change over time and respond to changes in environmental variables. Therefore, both baseline and range values were normalized by taking a ratio between the sensor estimates and reference measurements and expressed as percentage.

Drift was quantified by finding the difference between median sensor device-to-reference ratios for maximum and minimum quartiles of each factor (i.e. time, temperature, humidity, pollutant concentration). For example, the inter-quartile difference for change over time was found by calculating the difference in median sensor device-to-reference ratios between August and April of the following year.

The inter-quartile difference for temperature was found by calculating the difference in median sensor device-to-reference ratios for low temperature range (-13-8 C) and high temperature range (20-32 C) based on the variation measured over the selected 9-month period. The impact of humidity and pollutant concentration was quantified similarly.

In order to compare AirSENCE devices based on the extent to which reproducibility (variation between device replicates) changes over time, inter-device %RSD was analyzed for 4 device replicates. It was calculated by finding the standard deviation for a given pollutant across 4 devices and dividing it by the average for these devices (Figure 36).

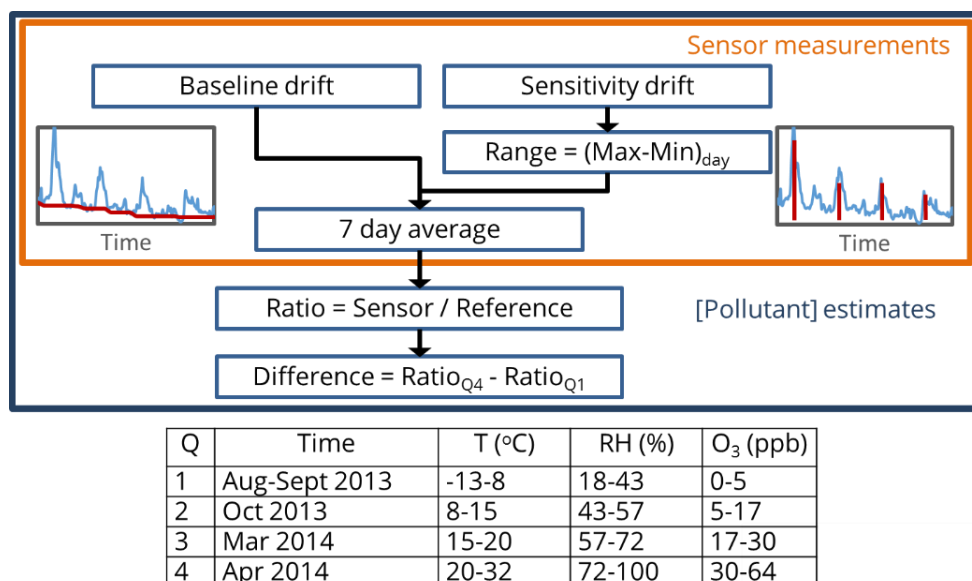


Figure 36. The process for analyzing the impact of drift and interferences on pollutant model performance.

4.3.1 Pollutant model drift analysis

The major goal of sensor device drift analysis was to evaluate whether there is any appreciable degradation in sensor performance and make projections about the sensor operation lifetime. Understanding how drift influences the sensor device performance helps identify methods for its correction. For example, drift in baseline can be corrected for by rescaling the sensor device performance or inter-comparisons between sites during the periods when pollutant levels are low. However, drift in sensitivity is more difficult to correct. More complicated and expensive measures such as device re-calibration or sensor replacement might be needed to correct for sensitivity drift. In addition, understanding of the impact of drift on inter-device variation as well as the causes for this change helps assess how the ability to resolve spatial variability in pollution varies over time.

The impact of drift on change in sensor baseline, sensitivity and inter-device variation was evaluated between August 2013 and April 2014. The change in inter-device variation over time was also compared for different pollutants.

Baseline and sensitivity were found to decline over time (Figure 37). The change in sensitivity (i.e. range) was bigger than that of baseline and was more consistent between device replicates. Unlike the other pollutants, the $PM_{2.5}$ baseline estimates increased over time, which was likely caused by accumulation of dust on the scattering optics. O_3 and $PM_{2.5}$ models were more impacted by the drift in baseline, while NO_x , CO and CO_2 estimates were more impacted by drift in response range (sensitivity).

Some devices were impacted by large change in sensitivity at the onset due to low quality of some sensors, suggesting that early diagnosis of sensor operation parameters is vital. However, the long-term drift in baseline and range was found to be consistent across device replicates and independent of seasonal variation in temperature, humidity and pollutant concentration. The response range declined to less than 60% of reference after 2 months of operation and continued to decline to below 30% after 4 months of operation. For most devices, drift correction or re-calibration is needed after 2 months of operation. After 6 months of operation sensor replacement might be needed for some sensors.

The variation between devices was also found to decrease over time for NO_x , CO, CO_2 , and to a lesser extent – O_3 and $PM_{2.5}$. While initially %RSD is close to average % reference residual, it decreases by over 50% in 9 months. The main reason for the decline in variability between devices over time was the decline in sensor range. As the sensors aged, key sensors became less responsive to changes in the pollutant concentration and thus their readings approached a constant value. With each sensor showing less variability, the variability between sensors also declined and hence the %RSD decreased. Based on the earlier analysis (Section 4.1.7), the key sensors responsible for pollutant model drift and high inter-device variation were P/N706 and PPD42.

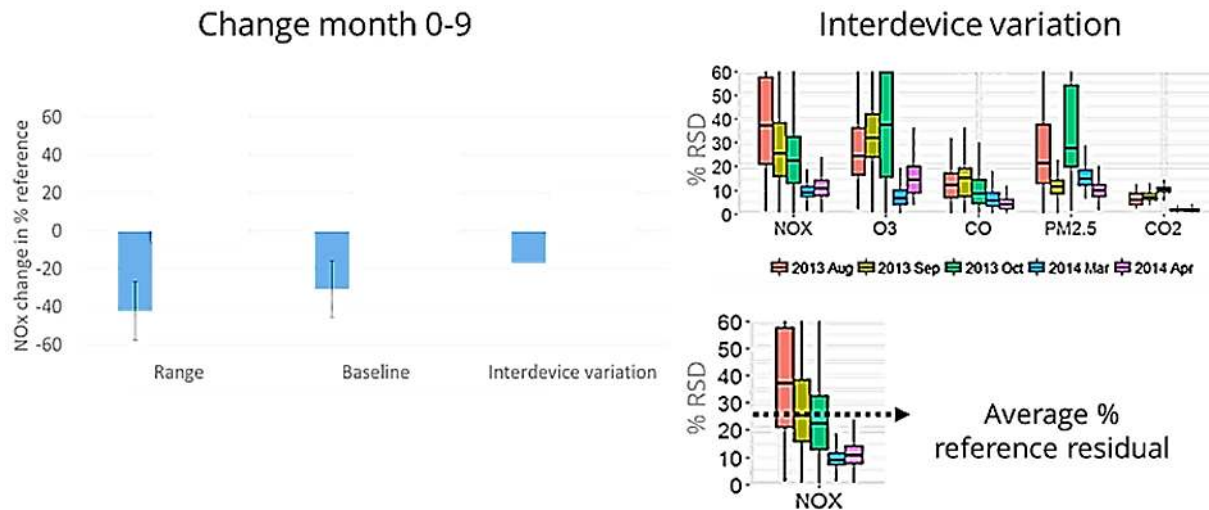


Figure 37. Impact of time on pollution baseline and range measured by AirSENCE devices. Impact is expressed as change over time (difference between end of Spring 2014 and Fall). Both baseline and range declined over time as indicated by negative % change in baseline and range. Average of 3 device replicates is shown. In addition, comparison of inter-device variation (%RSD) for NO_x, O₃, CO, PM_{2.5} and CO₂ is shown as a function of time. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

The change in sensor device performance over time also impacts the relationship between pollutants (Figure 38). While the slope between CO and NO_x measured by reference monitors doesn't change over time, the slope and correlation between CO and NO_x for AirSENCE devices increases. CO and NO_x sensors become more correlated and cross-responsive as baseline and range of sensor response declines and hence became less capable of resolving CO from NO_x.

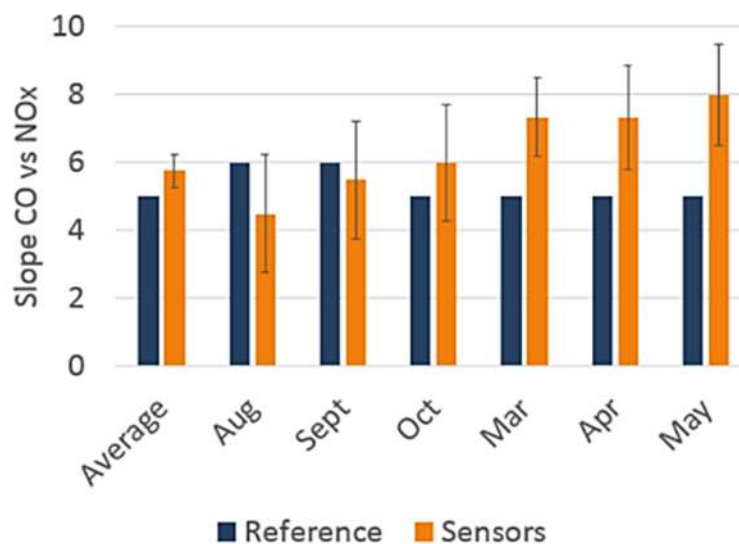


Figure 38. Drift results in inaccurate long-term assessment of multipollutant patterns. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

4.3.2 Interference impact

Combining different types of sensors contributes to more robust models that are less sensitive to the influence of some interferences. For example, combining O_3 and NO_x sensors in a NO_x calibration model helped correct for the corresponding interference. However, depending on the quality and variability of the training dataset used, the regression models were still found to be impacted by interferences.

The impact of cross-sensitivity was assessed by comparing the strength of correlation between sensor pollutant estimates and target pollutant concentration. For instance, the strength of correlation between the sensor NO_x estimates and the reference NO_x measurements indicates how good the devices are in predicting target pollutant levels, while the strength of correlation between the sensor NO_x estimates and the reference measurements for other pollutants (e.g. CO , CO_2 , O_3) indicates whether these pollutants might interfere with sensor response.

Interferences were found to have significant impact on model performance (Figure 39). The strength of the correlation between NO_x sensor device estimates and reference measurements for NO_x , CO_2 , O_3 and CO was used as a proxy to assess the performance of sensor device NO_x MLR model. The correlation was lower than 1 (0.7) for NO_x . The correlation was also smaller between sensor NO_x and

reference CO (compared to reference NO_x and reference CO). However, the correlation was higher between sensor NO_x and reference CO₂ and O₃ (compared to the correlation with reference NO_x). These results suggest that multi-sensor models can under- or overestimate the actual correlation between pollutants. One of the reasons for reduction in model accuracy and selectivity was the quality of the training dataset used to build the multi-sensor models. When the training dataset does not contain sufficient variation in environmental conditions and pollutant concentration range, the calibration models are not able to effectively account for the interferences. Thus, the cross-responsiveness of the sensors cannot be effectively corrected for. In addition, pollutants in ambient air were often more correlated during some periods of time compared to others, highlighting the importance of using sufficiently large and variable training datasets.

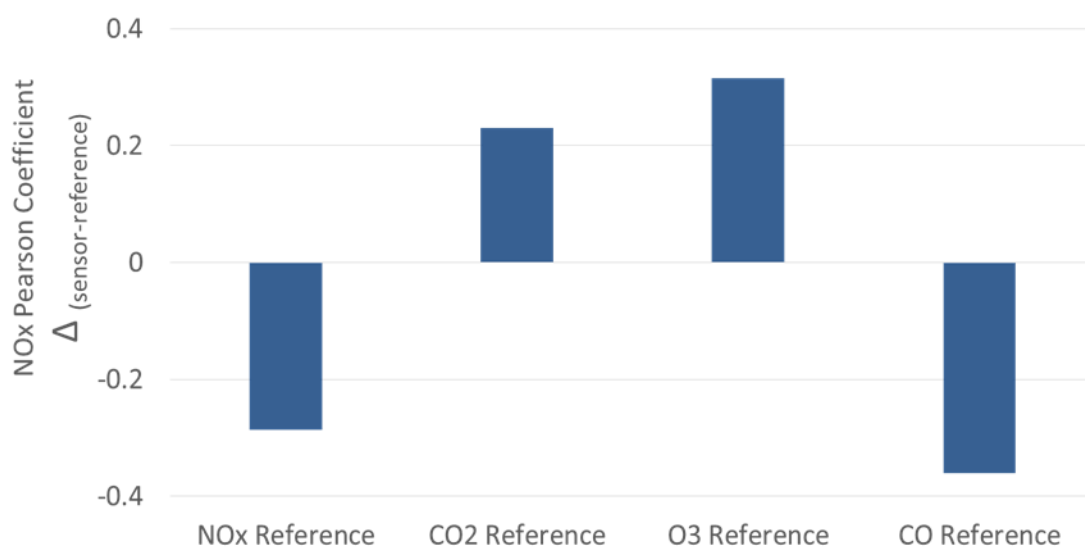


Figure 39. Impact of inferences on NO_x sensor response, expressed as the difference in strength of the correlation between sensor device output and reference instruments ($R^2_{\text{Pearson NOx sensor}} - R^2_{\text{Pearson NOx reference}}$). Pearson correlation coefficient was calculated between sensor device NO_x and reference measurements for NO_x, CO₂, O₃ and CO ($R^2_{\text{Pearson NOx sensor}}$). The extent of error in the correlation strength was then quantified by subtracting the Pearson correlation coefficient for reference NO_x and reference measurements for NO_x, CO₂, O₃ and CO ($R^2_{\text{Pearson NOx reference}}$). The value of $R^2_{\text{Pearson NOx reference}}$ for reference NO_x vs. reference measurements for NO_x was 1.0. Average of 6 AirSENCE devices is shown. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

4.3.3 *Influential factors*

The assessment of influential environmental factors on sensor response was discussed in Section 4.1.3. To extend that analysis and assess the extent to which MLR models can correct for influential factors was further investigated by comparing the impact of temperature, humidity, interfering gases on sensor model performance. The major goal of this analysis was to identify the most influential variables to enable better data quality control and interference correction. Insight of the factors influencing sensor response helps: 1) flag data for quality control and 2) develop more robust models by ensuring more variability in training dataset and improving the ability of AirSENCE devices to extrapolate beyond the training variable variation range. The impact of pollutant concentration on inter-device variation was also tested to assess the extent to which it impacts device performance and identify which pollutant models are most sensitive.

Ozone was found to have one of the largest impacts on inter-device variability of NO_x, O₃, CO and CO₂ (Figure 40). As the concentration of ozone increased, the variability between devices declined. While the inter-device %RSD for NO_x is close to average % reference residual at ozone concentration below 5ppb, it decreases with increasing ozone until it reaches 10% for ozone concentration higher than 17 ppb. In contrast, earlier analysis showed that inter-device sensor variability increased with pollutant concentration suggesting that the calibration models subdue the variability in sensor replicates at higher concentration of ozone. The decrease in inter-device variation with increased pollutant concentration might also be related to reduction in sensitivity of regression models towards changes in pollutant level at higher concentration. Drift over time also has a similar effect on inter-device variability suggesting that significant reduction in inter-device variability can be a sign of degraded performance.

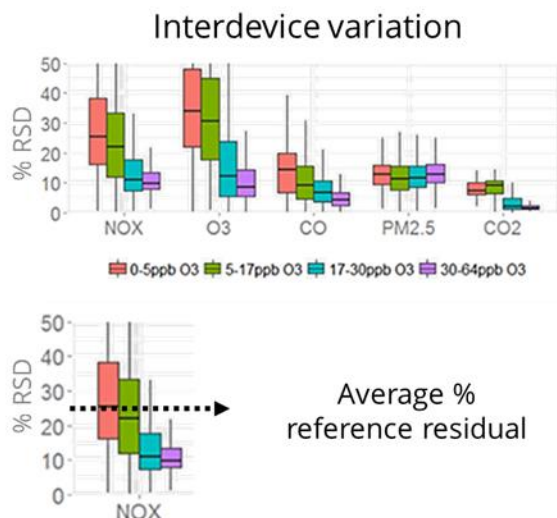


Figure 40. Impact of pollutant concentration on inter-device variation. Comparison of inter-device variation (%RSD) for NO_x, O₃, CO, PM_{2.5} and CO₂ is also shown as a function of pollutant concentration. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

Increase in temperature was found to affect the greatest number of sensors (Figure 21). Its impact on sensor device response range, baseline and inter-device variation was therefore considered in more detail. Its impact on inter-device variation was also compared for target pollutants. To correct for the influence of seasonal change in pollutant concentration, a sensor-reference ratio was used as a metric instead of pollutant concentration.

Temperature was found to have the largest impact on increasing regression model baseline and the increasing range, particularly for NO_x (Figure 41). Humidity also had an effect of increasing baseline and range, albeit much smaller one. The impact of temperature could be confounded by the impact of sensor drift since most of the measurements in low temperature conditions co-occurred with sensor aging (Section 4.3.1). However, even when the dataset was limited to the first 3 months of sensor operation, the temperature influence trend remains significant suggesting that the impact of temperature and drift with time are separate. The impact of temperature on sensor performance is also supported by the literature (Popoola et al. 2016). These results suggest that the device performance improves as temperature and humidity increase and that the control of environmental conditions within the sensor box could help improve device performance.

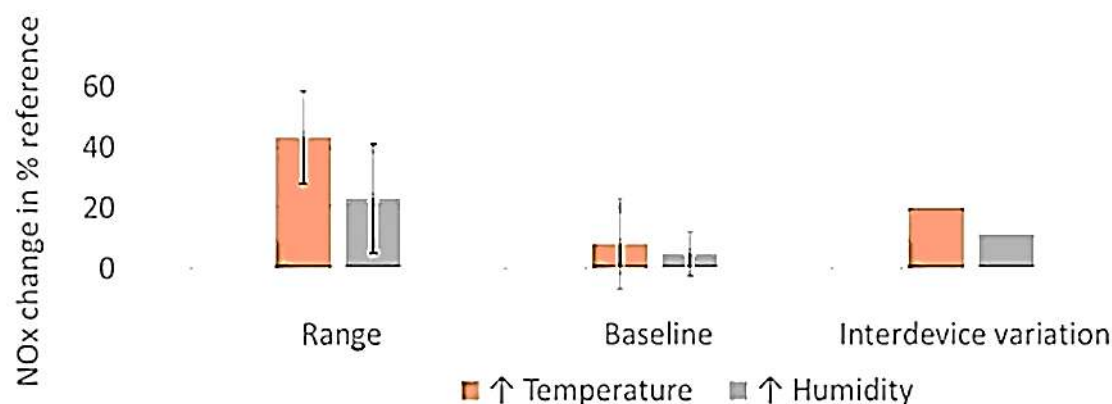


Figure 41. Impact of temperature and humidity on pollution baseline and range measured by AirSENCE devices. Impact is expressed as difference between lowest and highest quartiles. Positive change indicates increase with the factor. Average of 3 device replicates is shown. Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

4.3.4 Site variation

The influence of interferences on sensor device response suggested that sensor device calibration models might be sensitive to variation in pollutant mixtures in locations with different types of sources and levels of air pollution. In addition, earlier results showed that the importance of regression model predictors changes depending on the type of site. Therefore, the evaluation of performance of AirSENCE devices at sites with different levels of air pollution is important to understand the accuracy and limitations of spatial variation analysis.

Therefore, reproducibility of regression model response was evaluated for AirSENCE devices trained at SOCAAR site and deployed at the Toronto Island Halan's Point and Highway 401 sites. Toronto Island site is an example of background site with low level of traffic pollution, while Highway 401 site is located close to highway and is an example of site with high levels of traffic pollution. SOCAAR site is close to an arterial road and is affected by medium levels of traffic pollution.

The analysis resulted showed that good performance at calibration site is not indicative for acceptable performance at other sites (Figure 42). While AirSENCE devices showed acceptable performance when

trained and tested at SOCAAR, the device accuracy was degraded at the island and highway sites.

Regression models for NO_x trained at SOCAAR site can be used to distinguish medium and low pollution levels for NO_x but not levels above 40 ppb. The deterioration in model performance differs by the type of pollutant. SOCAAR-trained regression models fail at accurately measuring ozone at location with low levels of traffic pollution such as island site. They also cannot measure CO₂ and PM_{2.5} reliably at the sites where they were not trained. However, the performance of CO models trained at SOCAAR was shown to be acceptable and suitable for resolving sites with different levels of air pollution.

Two factors that have likely contributed to poor regression model performance are interferences and sensor surface saturation. Sensor response is correlated to exchange of oxidative and reductive species on sensor surface. Hence, the range of sensor response is limited and depends on kinetics of pollutant reactions on the sensor surface. In addition, high sensor cross-responsiveness suggests the importance of the training dataset. When calibration models are developed, they are optimized to correct for specific types and levels of interferences. The change of device location leads to changes in the types and levels of both target pollutants and interferences, which can lead to lower accuracy. Development of calibration models using data from multiple sites can lead to more robust performance and better ability to access spatial variability of air pollution.

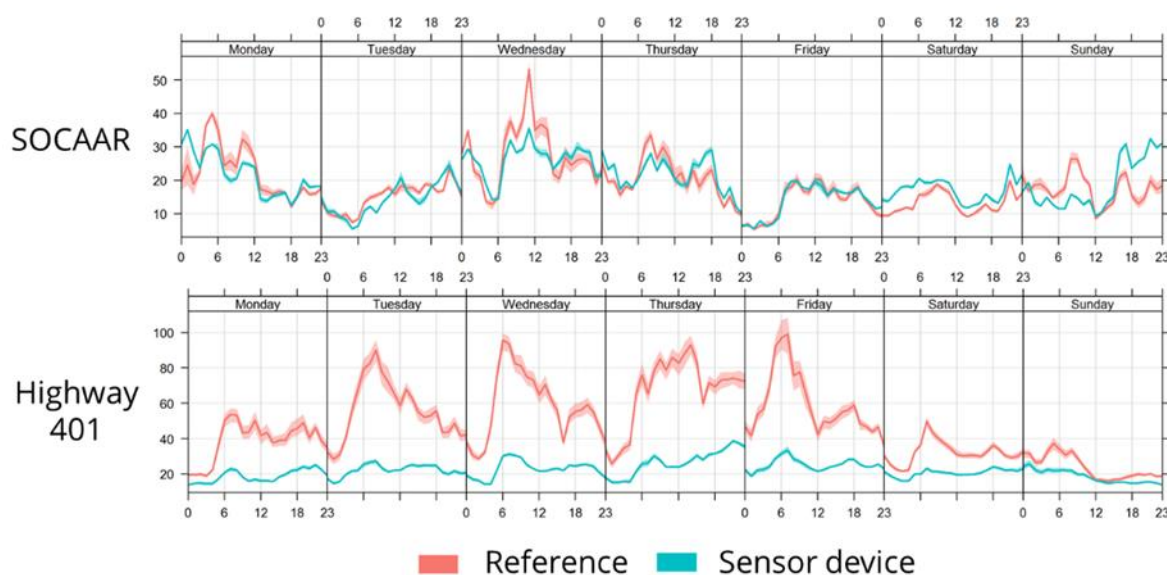


Figure 42. NO_x model performance before (top) and after (bottom) deployment at the highway site. Average weekly trend over a month is shown. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.3.5 *Summary*

Drift was found to have a large impact on sensor device model performance. The response range declined to less than 60% of reference after 2 months of operation and continued to decline to below 30% after 4 months of operation. The impact on sensor sensitivity was higher than on baseline response and affected slope and correlation of pollutant regression models. Inter-device variation declined over time for most pollutant models due to the progressive decline in the responsiveness of the devices.

Environmental factors were also found to have large impact on pollutant model performance. Increase in temperature and humidity caused increase in pollutant model baseline, sensitivity and inter-device variation. Temperature had bigger impact on model performance than humidity and the impact on pollutant model sensitivity was bigger than on baseline.

Gas interferences also had a large impact on regression model performance. NO_x regression model was found to be highly correlated with O₃ and CO₂ suggesting the impact of these gases on model performance. The effect of drift, environmental factors and interferences was highest for NO_x and O₃.

Good performance at a calibration site was also found to not necessarily assure acceptable performance at other sites. The accuracy of calibration models was found to decline when devices were deployed at new locations. The degradation in performance was smallest for CO and largest for CO₂ and PM_{2.5}. For NO_x and O₃, the performance was dependent of the range of pollution concentration at the site.

4.4 Performance optimization

While AirSENCE devices have acceptable performance when tested at the same site they are calibrated at, the accuracy declines when the devices are deployed at sites with different types of sources or pollution levels. The two major reasons for decline in performance are: 1) sensor non-linearity, 2) sensor cross-responsiveness. While linear regression models are preferable due to their simplicity and speed of computation, their accuracy is low for many pollutants due to nonlinearity of sensor response. Decline in accuracy is often caused by underestimation of pollutant concentration and is most significant at higher pollutant levels. The large impact of non-linearity on sensor device model performance suggests

that nonlinear models and variable transformations might improve the accuracy of the pollutant estimations.

Sensor cross-responsiveness is another inherent characteristic of commercial MOS sensors. While it can be partially corrected for using regression models, the performance of the models is often biased by the training dataset. When calibration models are developed using data from a single location with low variability in the types of sources and emissions, they are often limited in the types of factors and interferences they can correct for. As a result, deployment at a new site often leads to decreased accuracy because of change in types and combinations of pollutants. The quality and variability of the training dataset was shown to have a large impact on pollutant model accuracy at sites where pollutant levels are higher or lower compared to the model training site. Using data from multiple sites to calibrate AirSENCE devices might therefore improve device performance.

To correct for the impact of sensor nonlinearity and cross-responsiveness, several different techniques were tested. Nonlinear regression models such as support vector machines and random forests were compared to multiple linear regression (MLR) models in their ability to improve the accuracy of pollutant baseline and range estimates. In addition, the impact of different types of transformation on sensor response was assessed. The impact of variable transformation on MLR model performance and improvement in performance across sites was also tested. In addition, the impact of multi-site calibration on performance of calibration models and their ability to correct for sensor cross-responsiveness was also tested.

The following questions were investigated:

Nonlinear models impact

1. Can nonlinear models be used to improve sensor device performance? What is their impact on sensor baseline and range?

Variable transformation impact

2. What types of variable transformation techniques are most impactful in improving the device performance? How does variable importance ranking change when variables are transformed?
3. How do MLR models using transformed predictors compare to MLR models using raw sensor measurements when their performance is tested at sites with different level of traffic pollution?
4. How do MLR models developed using transformed variables compare in their ability to resolve special differences between sites?

Training data impact

5. How to AirSENCE devices calibrated at sites with low, moderate and high traffic pollution compare in their performance?

4.4.1 *Nonlinear models*

Linear regression makes a number of assumptions regarding the variable distribution as well as relationships between predictors:

- Linear relationship between independent and dependent variables
- Multivariate normality (normal distribution of predictors)
- No or little multicollinearity (low correlation between predictors)
- No auto-correlation (residuals are independent of each other)
- Homoscedasticity (error is the same across all values of predictor)

Earlier analysis showed that while majority of the assumptions were valid, some didn't apply across full pollution concentration range. The limitations in pollutant concentration range for which the assumptions applied became apparent when the devices were deployed at other sites. For example, due to sensor response non-linearity, the calibration models underestimated the pollutant levels at some sites and overestimated them at other sites. This suggests that the models were not generalizable and that alternative regression techniques might be needed to improve model performance.

Alternative regression techniques were tested to verify whether model performance could be improved. These techniques included GLM (generalizable linear model), MARS (multivariate adaptive regression spline), NN (neural network), SVM (support vector machine) and RF (random forest). Most of the selected alternative regression techniques were non-linear (except for GLM) to help validate the hypothesis that non-linearity in sensor response contributed to poor calibration model reproducibility. The analysis showed that nonlinear models improve accuracy for some pollutants (Figure 43). According to the comparison of model performance in lab and in field, SVM models were found to improve model predictions for PM_{2.5}. MARS, GLM, RF and NN models did not significantly improve the model predictions and in some cases even reduced the accuracy.

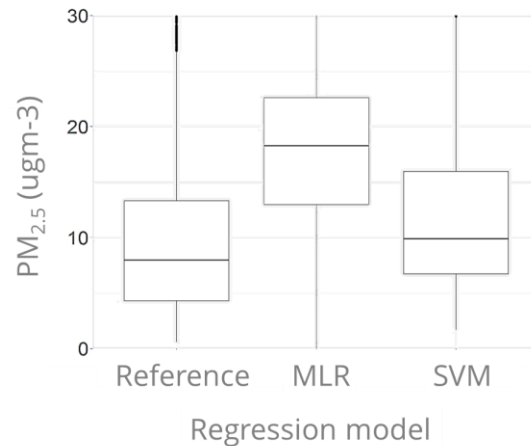


Figure 43. Analysis of improvement in performance when nonlinear calibration models are tested. Comparison of MLR and SVM models for monitoring $PM_{2.5}$ at the Highway 401 site is shown. All models were trained at SOCAAR location. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.4.2 Transformed sensor importance

Although nonlinear models improve the device performance, regression models like SVM and RF are computationally expensive when used in real-time calculations. Transforming sensor response prior to using them as predictors in MLR models is another method that can be used to correct for nonlinearity in sensor response, while reducing the computational needs (i.e. time and memory required to generate real-time model estimates).

Variable transforms were investigated as a method to correct for sensor measurement non-linearity and improve the model accuracy. The impact of variable transforms on model importance was evaluated by:

1. Calculating variable transforms for all sensor variables
2. Selecting the variable combinations that had the highest performance
3. Evaluating the improvement in accuracy for AirSENCE devices trained at SOCAAR site and deployed at other sites.

The following transforms were calculated for each sensor variable for all devices:

- Natural logarithm transform
- Base 10 logarithm transform

- Exponential transform
- Square transform
- Square root transform

Relative importance metrics for all variables were calculated. SOCAAR data and data from other sites (i.e. Toronto Island and Highway 401) was used for model training and validation respectively. The relaimpo R package was used to compute variable importance and adjusted R squared was used as a metric. New predictor sets were chosen based on variable importance analysis.

The analysis has indicated that square and exponential transform were the most useful in improving the model performance (Figure 44). In addition, sensors TGS880 (important in NO_x, CO, CO₂ models) and MICS5521 (important in NO_x and O₃ models) were found to be more influential on model performance when transformed.

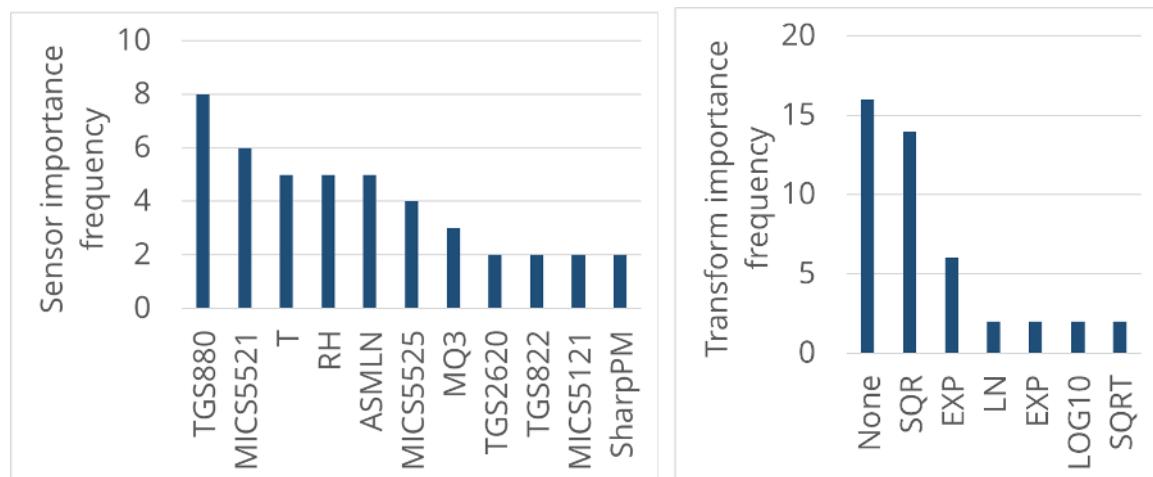


Figure 44. Best-performing sensor transforms.

4.4.3 Variable transformation impact

While nonlinear regression techniques and variable transforms contributed to improvement in model performance, the AirSENCE devices still underestimated the peak pollutant levels, particularly for sites with high traffic-associated pollution. A hypothesis explaining the reason behind peak pollutant level underestimation was regression model bias caused by skewed pollutant concentration distribution. The occurrence of high pollutant concentrations was relatively infrequent. As a result, most of the training data was much lower than the peak pollutant levels, which could bias the calibration models. This bias

was further amplified since the calibration models were developed at SOCAAR site, where peak pollution levels are lower compared to higher traffic locations, such as Highway 401 site.

The influence of peaks and baseline on calibration model was therefore decoupled by building calibration model using background-subtracted data. Background was calculated by finding the rolling minimum average with a 60-minute window for both reference data and sensor data (Figure 45). Background was then subtracted from the raw data for both reference instruments and sensor measurements to find the local source contribution to the measurements. Two separate models were then built: 1) background regression model and 2) peak regression model. The models were then applied to new sensor data and added to generate pollutant concentration estimates.

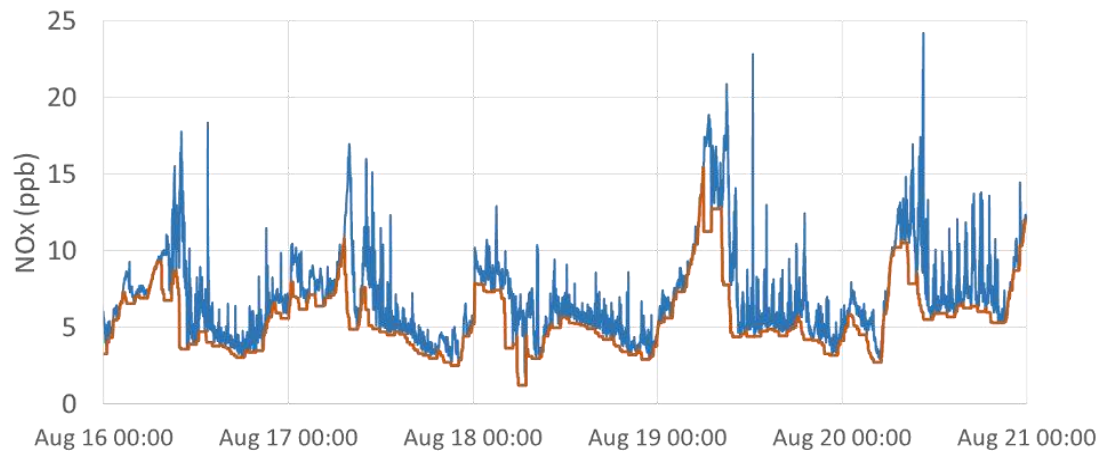


Figure 45. Baseline subtraction for NO_x measurements at site 18. Baseline is indicated in orange. One-minute resolution data is shown.

It was found that the combination variable transformation and separate model development for local and background components together led to significant improvements in regression model accuracy for NO_x, O₃, CO and CO₂ at both Highway 401 and Toronto Island sites (Figure 46). The improvement in model accuracy for ozone was especially significant at the Toronto Island site. The accuracy of estimating both the median concentration as well as the concentration distribution improved.

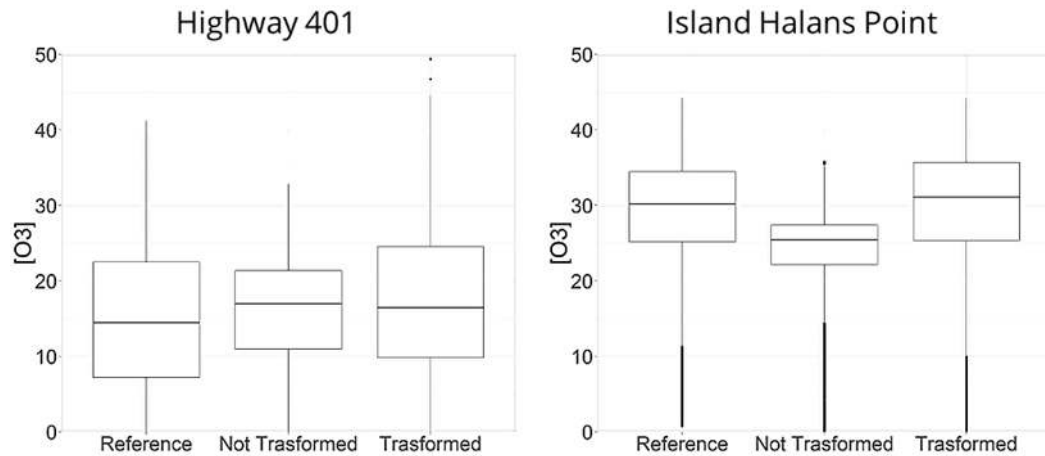


Figure 46. Comparison of non-transformed and transformed O_3 regression model performance tested at the island and highway sites. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.4.4 Training site impact

For most calibration models described in this thesis, 2-3-week period at SOCAAR was used for model training. While this period of time proved sufficient for development calibration models that performed well for extended periods of time when deployed at SOCAAR, the performance at other sites was significantly worse. One of the biggest challenges with MLR models was that for some pollutant targets they did not perform well at sites they were not calibrated at. It was hypothesized that training sensor device models using sufficiently variable data is important due to high cross-responsiveness in sensor response. Since sensors have been shown to be responsive to multiple pollutants and environmental factors, capturing more variation in the levels of pollutant concentration, temperature and humidity in the model training data could improve the model accuracy and reproducibility at other sites.

The impact of the calibration location of AirSENCE devices on model performance was evaluated. The AirSENCE devices were first trained at SOCAAR and their performance was verified at highway and island sites by comparing the similarity of concentration distributions between sites. The importance of calibration site was evaluated by training devices at highway or island location and testing the performance at SOCAAR.

It was found that calibration at site with higher pollutant concentration improves overall calibration model accuracy (Figure 47). For CO, the AirSENCE devices overestimated levels of CO concentration (by 50 ppb on average) at highway when the models were trained at SOCAAR. When devices were trained at the highway site, the overestimation of CO levels at highway was smaller (10 ppb) and the devices performed well when the model was applied to SOCAAR data. When devices were trained at SOCAAR and deployed at the Island, they performed well with regards to evaluating CO distribution. However, when the devices were trained at the Island and tested at SOCAAR, they were not able to correctly extrapolate the higher CO levels. This suggests that the CO calibration model cannot accurately extrapolate high CO levels outside of the calibration range. A major reason for this is nonlinearity in CO sensor response at higher CO levels. Training the CO model at a site with higher CO concentration levels is essential for accurate model output.

NO_x calibration models trained at SOCAAR site underestimated the NO_x levels (10 ppb on average) when deployed at the highway site. The models were able to correctly assess the NO_x concentration when trained with data from the highway site but were unable to extrapolate the lower NO_x concentrations correctly. This suggests that the NO_x regression model is nonlinear at both higher and lower concentrations and requires training at multiple sites. In addition, nonlinear regression techniques or variable transformation might assist in more accurate model output.

These findings support the hypothesis that calibration of AirSENCE devices using skewed data can bias pollution estimates, especially when peak pollution levels at the calibration site are lower than those at deployment site. The results suggest that either multi-site calibration is key to accurate analysis. Another solution is a strategic calibration protocol where devices are calibrated at a site with higher pollutant levels than the site where they are deployed.

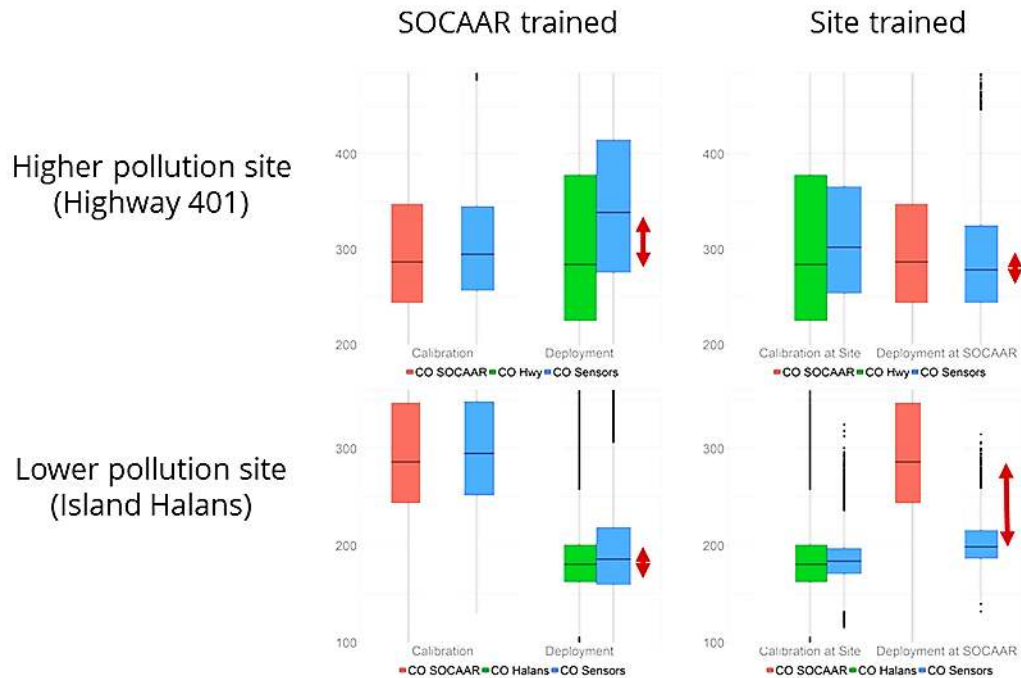


Figure 47. Comparison of CO calibration model trained at SOCAAR and tested at Highway 401 and Island sites with CO regression model trained at the sites and tested at SOCAAR. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

Another key finding was that using more than 2 months of historical data didn't improve the model performance. Training the AirSENCE devices at a specific site helps generate best calibration models at that site, but training AirSENCE devices at multiple sites leads to the best performance across multiple sites.

4.4.5 Summary

In summary, nonlinear models were found to improve the accuracy of regression models for NO_x and PM_{2.5}. However, continuous deployment of these models was computationally expensive. While variable transformation was found to improve the MLR model accuracy, AirSENCE devices have a tendency to underestimate the pollutant levels, particularly at sites with high pollutant levels. The combination of variable transforms with separate model development for background and local contributions helped improve model performance, particularly at sites other than the calibration site.

In addition, the reproducibility of models at other sites was improved when devices were calibrated at sites with higher pollutant concentration. The sensor device model performance and ability to resolve sites was highest when the devices were calibrated using data from several different sites. A major reason for improved performance was the improved ability of calibration models to adjust for different combinations of interferences. However, long-term historical training data didn't significantly improve model performance.

4.5 Temporal patterns

The ability to track the variation of air pollution over time is useful for understanding air pollution sources. The amplitude and integral of short-term spikes in pollution levels can help appreciate local sources such as traffic. The variation of pollution levels at time resolution less than 1 minute can be used for analyzing traffic emission factors. Long-term temporal analysis can be used to assess the changes in impact of source profile and meteorological variables of pollutant levels. The analysis of the absolute pollutant levels by AirSENCE devices is often confounded by sensor nonlinearity and cross-sensitivity to interferences. However, even when the absolute pollutant estimation is not accurate, the relative temporal patterns are still informative in comparing the changes in sources contributions. In addition, earlier sensor device validation analysis indicates that the devices are able to assess temporal patterns better than absolute pollutant levels.

The performance of AirSENCE devices in detecting temporal pollution patterns was assessed by comparing diurnal variation at different sites. Both background and local contributions were considered and the differences in peak frequency and amplitude between sites were investigated for sites with different levels of traffic. The variation between sites was validated through: 1) comparison with known temporal variation in pollutant concentration (e.g. morning rush hour) and 2) comparison with known traffic volume differences between sites. The sensor device performance in detecting diurnal, weekly and seasonal patterns was validated against reference monitor at SOCAAR, high traffic site and background site. The temporal variation in relative contributions of gasoline and diesel sources was also inferred from temporal variation of CO/NO_x ratios.

The following questions were addressed:

Temporal pattern estimation accuracy

1. How accurate are the AirSENCE devices in detecting temporal variation of pollutants over the course of a day, a week and a year?

Spatial variation of temporal patterns

2. Can AirSENCE devices be used to detect differences in pollutant diurnal patterns at different types of sites?
3. How does the accuracy of temporal patterns measured by AirSENCE devices vary between sites with different levels of traffic pollution?

Source contribution analysis

4. Can temporal patterns measured by AirSENCE devices be used to distinguish the sites with different levels of traffic pollution?
5. How accurate are the AirSENCE devices in assessing the temporal variation of relative contribution of different pollutants (and hence sources) at a site?

4.5.1 *Temporal pattern estimation accuracy*

A key application of AirSENCE devices is the analysis of changes in pollutant concentration over time. To ensure reliable analysis, the AirSENCE devices need to fulfill several criteria:

1. Accurately detect variation in pollutant baseline over time
2. Detect full scale of changes in pollutant levels
3. Show reliable performance in response to changes in pollutant levels and environmental parameters over time

While fulfillment of all the criteria is ideal, it is not always required for deriving insights from sensor device data. Even when the absolute magnitude of pollutant variation is inaccurate, the timing of the changes in pollutant levels can be useful for detecting the variability in source profile. Therefore, it was important to quantify the accuracy and reliability of AirSENCE devices in detecting the temporal variation in pollutant levels.

The accuracy of temporal patterns estimated by comparing the temporal variation in sensor measurements to that of reference monitors. The reliability in estimating the temporal patterns was evaluated by summarizing the accuracy over 5-month period. The temporal variation in sensor measurements was compared to reference monitors over a course of a day, a week and a year (Figure 48). Daily variation analysis was used to assess the ability of AirSENCE devices to detect changes in pollution levels due to rush hour. Weekly variation analysis was used to assess the reliability of AirSENCE

devices in response to small changes in source contribution. Assessments of temporal variation over the course of several months was used to understand long term reliability of sensor device performance and quantify the ability of AirSENCE devices to resolve seasonal patterns. The temporal variation was examined as the change in the average distribution 1) over the course of a day, 2) between weekdays and weekends as well as 3) over the period of 5 months.

The analysis has indicated that both short term and long-term temporal variation in pollutant levels can be resolved by AirSENCE devices. The diurnal variation in CO₂ levels due to rush hour was accurately and reliably detected by AirSENCE devices, although they slightly underestimated the peak levels in the early morning and over-estimated the reduction in CO₂ levels in late morning and early afternoon. The distribution median was also systematically higher for AirSENCE devices. The difference in pollutant levels between weekday and weekend was small for both reference monitors and AirSENCE devices. However, the distribution of CO₂ levels measured by AirSENCE devices was narrower than that of reference monitors. The AirSENCE devices were also able to accurately access changes in pollutant between May and September, although they tended to underestimate the pollutant levels in May and overestimate the levels in June. The accuracy in CO₂ estimates was highest in July, August and September when pollutant levels were lowest.

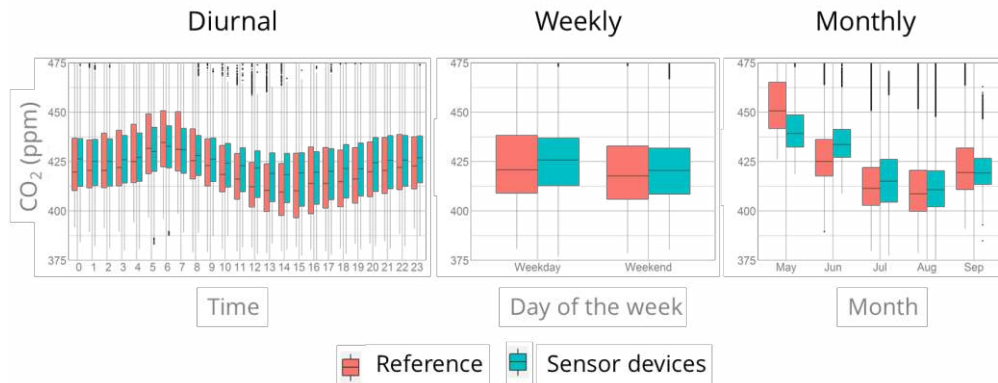


Figure 48. Evaluation of the ability of AirSENCE devices to resolve short-term and long-term variation in pollutant levels. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.5.2 Spatial variation of temporal patterns

The analysis of temporal patterns at a single site can help understand the variation in pollution sources and sinks. However, the comparison of temporal patterns between sites can be useful in detecting differences between sites even when the accuracy in detection of absolute pollution levels is insufficient. Temporal pattern parameters include:

- frequency of peaks
- amplitude of peaks
- timing of peaks
- magnitude of variation in baseline
- timing of variation in baseline

Time series analysis is a simple and intuitive way to assess how pollution levels and source contribution vary over both time and space. For example, the higher frequency of peaks may suggest higher volume of traffic, while the presence of high amplitude peak may suggest the presence of a source that emits a large amount of pollutant over a short period of time.

Diurnal temporal variation in pollutant levels was assessed at 5 different sites spread across Toronto during SCULPT campaign. The sites varied in the urban density, proximity to the road, size of local roads as well as the average volume of local traffic. Diurnal temporal patterns at the sites were compared with respect to magnitude of increase in NO_x levels during the morning rush hour.

It was found that both the magnitude and temporal pattern of the NO_x levels varied across the sites, suggesting that sensors can detect variation in temporal patterns (Figure 49). While large daily variation occurred at some sites (e.g. central site, site 12, site 18), the variation was much smaller at other sites (e.g. site 10, site 15). Morning period was important for understanding spatial variability. Most sites were characterized by an increase in NO_x in the morning, driven primarily by morning rush hour, but strength of morning peak differed between sites suggesting that sensors can detect variation in temporal patterns.

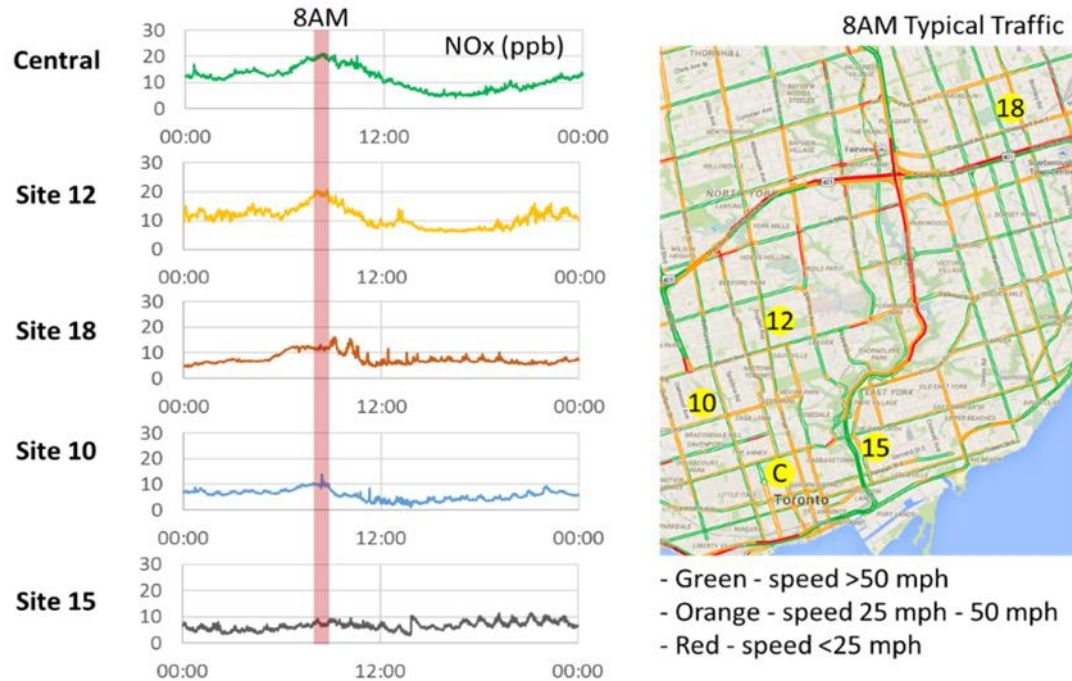


Figure 49. Comparison of NO_x temporal variation for AirSENCE devices at sites 12, 18, 10, 15, 5 and central site (1 min average data). Average of 4 weekdays is shown. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

AirSENCE devices were able to resolve the differences in temporal patterns between sites. However, the ability to detect differences between sites does not necessarily suggest accuracy. Factors such as local variations in interferences, temperature and humidity can affect the ability of AirSENCE devices to accurately resolve temporal patterns and some of the observed differences in temporal patterns could be due to interferences, not just changes in concentration. The validation of temporal variation of sensor device response at sites with different local source contributions is important for quantifying the confidence and limitations of temporal pattern analysis.

The validity of the observed temporal variation in sensor device response was assessed by comparing sites with low, medium and high levels of traffic. The devices were deployed at SOCAAR, Toronto Island and Highway 401 MOE stations between December 2015 and February 2016. The presence of reference pollution monitoring equipment at these sites allowed us to evaluate the sensor device performance. The sensor device pollution estimates were compared to the reference monitor by examining the temporal variation in the average pollution distributions over a course of day and over a period of 3 months.

The AirSENCE devices were found to accurately assess both the long-term and the short-term changes in O_3 levels at SOCAAR, where the devices were calibrated (Figure 50). However, even though they were able to resolve changes in O_3 levels with time, the AirSENCE devices failed to assess the full range of O_3 at high traffic site and also overestimated the median pollutant concentration. The AirSENCE devices were unable to reproduce the temporal variation at the low traffic site, even though the median pollution levels measured by AirSENCE devices and reference monitors were similar. The analysis suggests that while the temporal patterns and average concentrations estimated by AirSENCE devices might be different across sites, they are not necessarily accurate. Accuracy of pollution spatial variability assessment was also found to vary over time. Model optimization through calibration of devices using the data from multiple sites might improve the accuracy of temporal patterns estimated by AirSENCE devices.

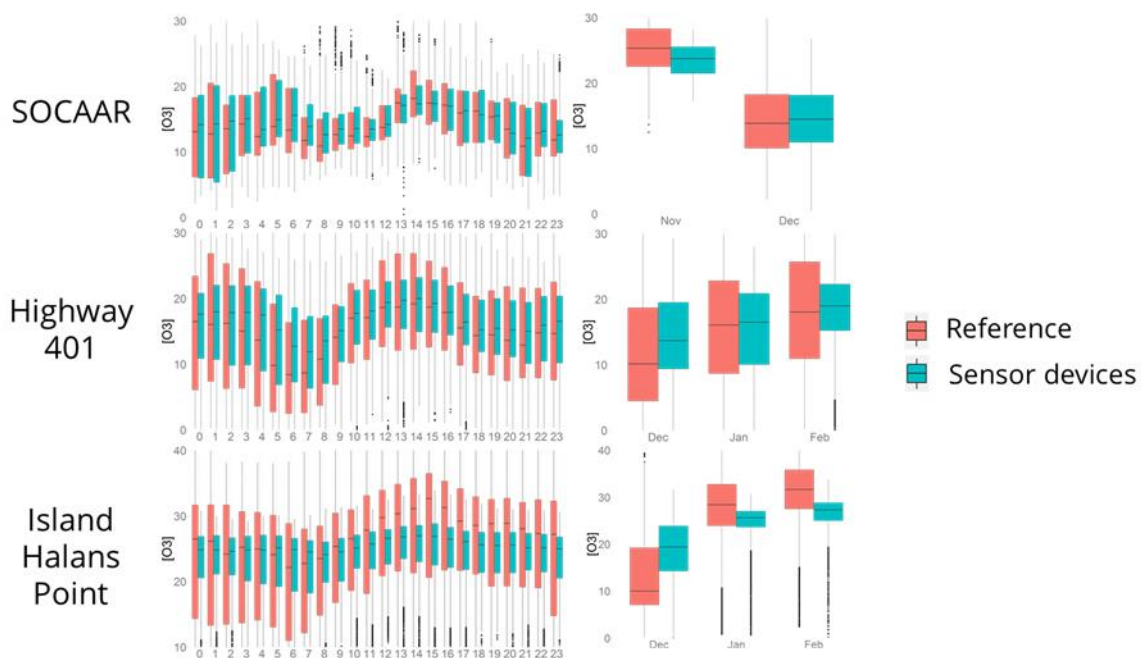


Figure 50. Evaluation of the ability of AirSENCE devices to compare temporal patterns of O_3 at 3 different sites. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.5.3 Source contribution analysis

In addition to site comparison, another key application of temporal variation analysis is the quantification of source emissions. The amplitude and area under the curve can be used for calculating emission factors and their variation over time. In addition, both the frequency and amplitude of peaks are related to the local traffic volume.

Analysis of the temporal variation of NO_x at sites with different levels of traffic was conducted to verify the hypothesis that the temporal variation in sensor device response can be used to rank the sites based on the level of traffic pollution. Each site in the SCULPT campaign was assigned an average level of traffic, measured in volume of vehicles per day within a 1500m buffer. Differences in NO_x temporal patterns were cross-referenced with the traffic volume labels.

The average NO_x levels at central site were found to be higher on average compared to sites with lower level of local traffic sources (i.e. site 18), particularly during the morning rush hour (Figure 51). In addition, the peak frequency was higher and the peak amplitude was lower for AirSENCE devices deployed at SOCAAR compared to those at the lower traffic sites, suggesting that while the traffic emissions might be contributing to more frequent changes in pollutant levels at SOCAAR, the magnitude of emissions from the sources at site 18 is higher. Site 18 is therefore likely influenced by a different type of pollution source, which produces higher emissions over a longer period of time.

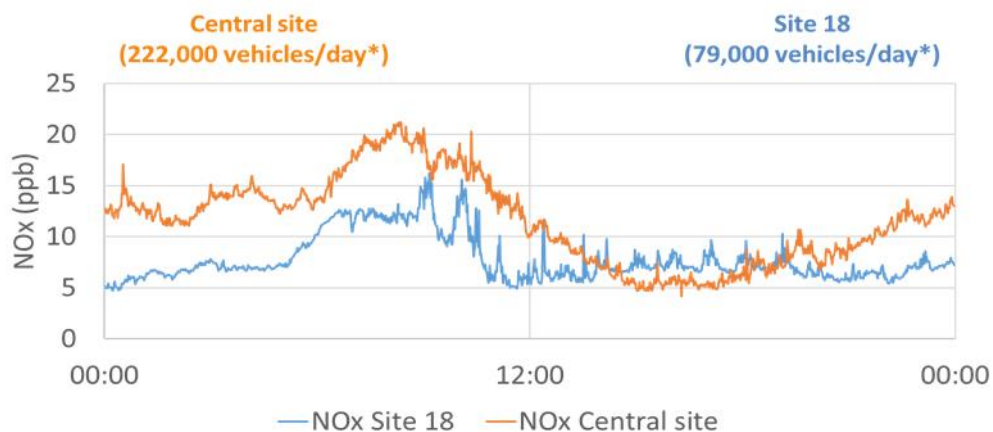


Figure 51. Comparison in differences in diurnal pattern for NO_x at site 18. Traffic level was measured in volume of vehicles per day within a 1500m buffer. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

The analysis of the variation of a pollutant over time gives insight into the magnitude and timing of the influence of local sources but doesn't reveal much information about the source category. The analysis of co-occurrence of peaks of different types of pollutants is helpful for attributing the pollution peaks to specific categories of sources. For instance, the relative amounts of NO_x and CO can help explain the relative contribution of gasoline and diesel sources. The relationship between these pollutants is known to vary between sites and over the course of a week.

To evaluate the ability of AirSENCE devices to distinguish the temporal patterns of NO_x and CO, CO/NO_x ratios were calculated. The average CO/NO_x ratios and distributions were compared for AirSENCE devices and reference instruments at the central site in order to establish the accuracy of sensor estimates.

AirSENCE devices were found to correctly detect changes in the CO/NO_x ratio over the course of a day, a week and in long term (Figure 52). However, the CO/NO_x ratios were underestimated at night and on the weekend, suggesting that comparison of gasoline and diesel sources might be less accurate on the weekend. There are two possible reasons for underestimation of CO/NO_x ratio: 1) underestimation of CO and 2) overestimation of NO_x. Earlier analysis showed that NO_x calibration models show lower correlation to CO and higher correlation to CO₂ and O₃ compared to reference instruments. Therefore, as the sources of NO_x and CO decline at night, the impact of interferences becomes more pronounced and this impacts the CO/NO_x ratio.

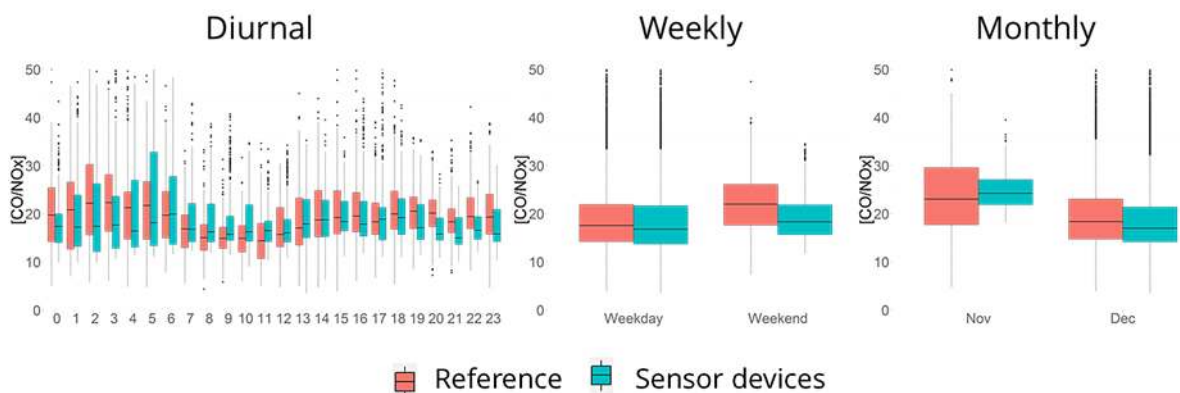


Figure 52. Evaluation of the ability of AirSENCE devices to accurately assess temporal variability of CO relative to NO_x. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.5.4 Summary

In summary, the analysis indicated that site difference in temporal variation can be detected. Both short-term changes in pollutant levels over the course of the day and long-term changes between months were detected.

Traffic-related pollution (TRAP) levels were highest and most divergent across most sites in the morning as a result of morning rush hour. The magnitude of morning increase in pollution levels measured by AirSENCE devices was found to vary between sites suggesting that sensors can detect variation in temporal patterns. TRAP levels produced by local sources were highest at the central site and site 12, medium at central site and site 18, and lowest at sites 10 and 15. The device performance was also validated against reference monitor data at sites with high, medium and low levels of traffic pollution. The accuracy of temporal pattern analysis was good at sites with high and medium traffic, but inadequate at sites with low traffic. There is evidence to believe that higher variability in the calibration model training dataset would lead to better performance.

Analysis of temporal patterns of NO_x revealed differences in peak magnitude, timing and frequency at different sites. The absolute pollutant levels were found to be related to the average traffic volume at the site. In addition, the timing and frequency of peaks revealed insights into local pollution sources. The AirSENCE devices were also found to detect changes in CO relative to NO_x, suggesting that they could be useful for comparing changes in relative contributions of diesel, gasoline and other types of combustion sources over time.

4.6 Spatial patterns

The ability to detect differences in pollution levels across sites is important for quantifying the population exposure to air pollution. However, the analysis of the absolute concentration is often confounded by sensor nonlinearity and cross-sensitivity to interferences. Comparisons of changes in pollution levels through analysis of temporal patterns can sometimes be used to rank sites based on the level of air pollution. However, the accurate assessment of absolute pollution levels is often required for the comparison of true contribution of local sources to pollution levels at different sites. It was important to identify the pollutant models that are more accurate in resolving differences between sites and assess what techniques could be used to improve the ability to detect spatial differences between

sites. In addition, analysis of the factors influencing the differences between sites contributes to understanding underlying causes. It was also important to understand which pollutants are more reliable for comparing differences between sites.

The performance of AirSENCE devices in detecting spatial pollution patterns was assessed by comparing the concentration distributions estimated by AirSENCE devices deployed at different sites. Pollutant models were assessed on their ability to distinguish sites with low, medium and high levels of traffic pollution. Methods such as predictor transformation were evaluated on their impact on the accuracy of detecting differences between sites. In addition, the change in the similarity between sites in response to different types of factors (i.e. time of day, day of the week and wind direction) was investigated.

The following questions were addressed:

Spatial pattern analysis accuracy

1. Can AirSENCE devices detect the spatial variation of absolute pollution levels?
2. Which pollutant models are the most accurate in detecting differences in pollutant distributions at sites with different levels of traffic pollution?
3. Can predictor transformation be used to improve the accuracy of MLR models in assessing the differences in pollutant distributions at sites with different levels of traffic pollution?

Spatial variation of temporal patterns

4. Does time of day affect the differences between sites? When are the sites most divergent from each other?
5. How do the differences between sites compare between weekdays and weekends?

Impact of wind direction on the differences between sites

6. What is the impact of wind direction on the differences between sites?

4.6.1 Absolute spatial pollutant variation

Traffic-related pollutants are known to be highly variable spatially. However, the assessment of their spatial variability is often limited because of small number of continuous air pollution monitoring stations available in most cities. For example, continuous NO_x measurements are only available at 4 Ministry of Environment stations in Toronto. The analysis of the magnitude and spatial scale over which the pollutant levels vary in the city is important for understanding the sources and their impact on population exposure. The relatively low cost of AirSENCE devices makes it possible to deploy them in

dense urban networks. However, the limited reference data makes it difficult to validate the differences between sites measured by AirSENCE devices. Land use regression models can be used in such cases to evaluate sensor measurements. While LUR models normally represent a snapshot in time, they can still be useful for comparing the average pollution levels in different regions.

An average NO₂ levels based on two LUR models was therefore used to assess whether AirSENCE devices can be used to detect spatial variation in absolute pollution levels of NO_x. The LUR models were calibrated from 2002 fall and 2004 spring monitoring campaigns with 143 monitors¹⁰. LUR model results were compared to sensor device data from the 5 different sites in Toronto that were studied during 2013 fall SCULPT campaign. The sites varied in the urban density, proximity to the road, size of local roads as well as the average volume of local traffic. While NO_x levels measured by AirSENCE devices are higher than NO₂ levels estimated by LUR models, they are proportional to each other and are still useful for site comparisons.

The detected difference in magnitude of pollutant levels the sites with high and medium concentration for AirSENCE devices during 2013 SCULPT campaign was found to be similar to the difference between the same regions as indicated by Land Use Regression (LUR) models (Figure 53). For example, the LUR-derived NO₂ levels in downtown were found to be 1.5 times greater than those in north-east region of Toronto. Similarly, the median NO_x levels measured by AirSENCE devices at downtown SOCAAR site were found 1.52 greater than those measured at site 18, in the north-east region of Toronto. Both median and distribution range of pollution levels were lower at site 18 compared to SOCAAR site. However, there were more outliers in NO_x levels at site 18, suggesting that while there are less local sources – the existing sources emit large amounts of pollution. Earlier temporal analysis has pointed towards similar conclusion. These results suggest that that AirSENCE devices are able to resolve absolute differences between sites both in terms of average pollutant levels and the peak levels.

¹⁰ Environmental Health Perspectives Jerrett et al. (2009)

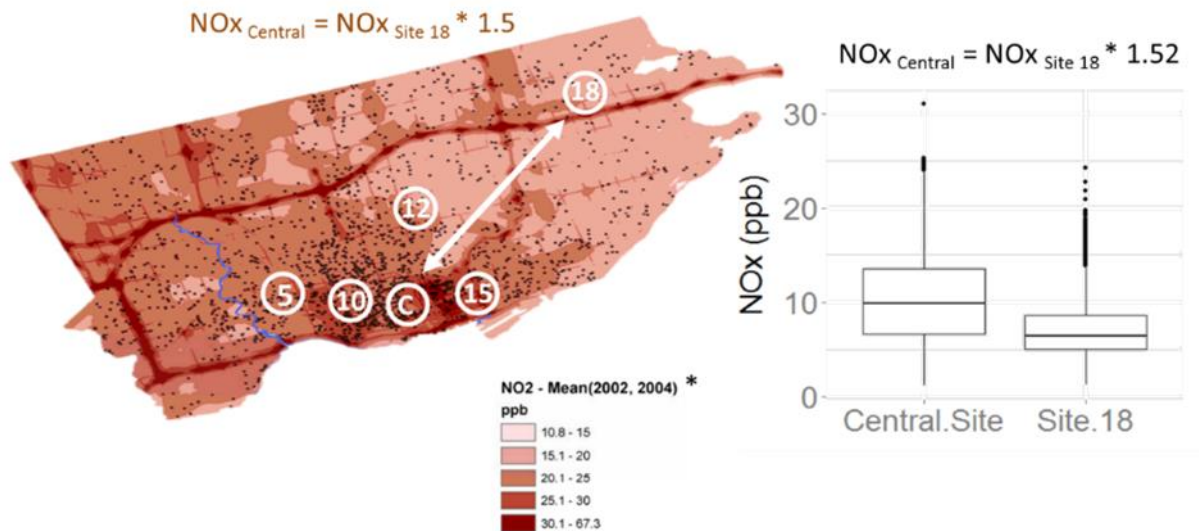


Figure 53. Comparison of the absolute NOx level differences measured by AirSENCE devices (shown in box-and-whisker distribution plots) and those derived using land use regression models calibrated from 2002 fall and 2004 spring monitoring campaigns with 143 reference monitors (shown in a map). Sensor data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

4.6.2 Validation of differences between sites

Model data suggests that AirSENCE devices show promise in resolving the differences between sites. However, it is important to validate the accuracy and limitations of spatial pattern analysis. Sensor device accuracy is known to be lower at sites other than the calibration site. Therefore, it was important to evaluate how accurate the sensor device models are in detecting differences in pollutant distributions at sites with different levels of traffic pollution. The goal was to identify the target pollutants for which the models are more reliable across sites and develop methods for correction of the models that are less reliable.

The validation of accuracy of sensor device was carried out at sites with low, medium and high level of traffic. The devices were deployed at SOCAAR, Toronto Island and Highway 401 MOE stations between December 2015 and February 2016. The presence of reference pollution monitoring equipment at these sites allowed us to validate the sensor device performance. Sensor device pollution estimates were compared to the reference monitor by examining the average pollution distributions at the three sites. All the sensor device models were calibrated at SOCAAR site.

CO model shows the best performance in detecting differences between low, medium and high traffic sites. Both distribution quartiles and median for CO estimates by AirSENCE devices were close to that of reference CO measurements (Figure 54). A low traffic site (i.e. Toronto Island) could be easily distinguished from the Highway 401 and SOCAAR sites because of lower mean and interquartile range. The high traffic site (i.e. Highway 401) has a median that is close to that of SOCAAR, but higher maximum values for both the sensor device and reference monitor. While the accuracy of the CO sensor device was highest compared to other pollutants, the AirSENCE devices still overestimated the 25% quartile for the highway and SOCAAR sites and 75% quartile – for the island site.

On the other hand, NO_x and CO₂ estimates by the AirSENCE devices could only be resolved at the SOCAAR and Island site because of underestimation of concentration at the Highway 401 site. O₃ estimates by the AirSENCE devices could only be used to compare the SOCAAR and 401 sites. The O₃ sensor estimates can be biased at sites with lower traffic pollutant levels because the calibration models have built-in correction factor for NO_x, which biases the O₃ estimates. PM_{2.5} measurements were not accurate enough to compare sites.

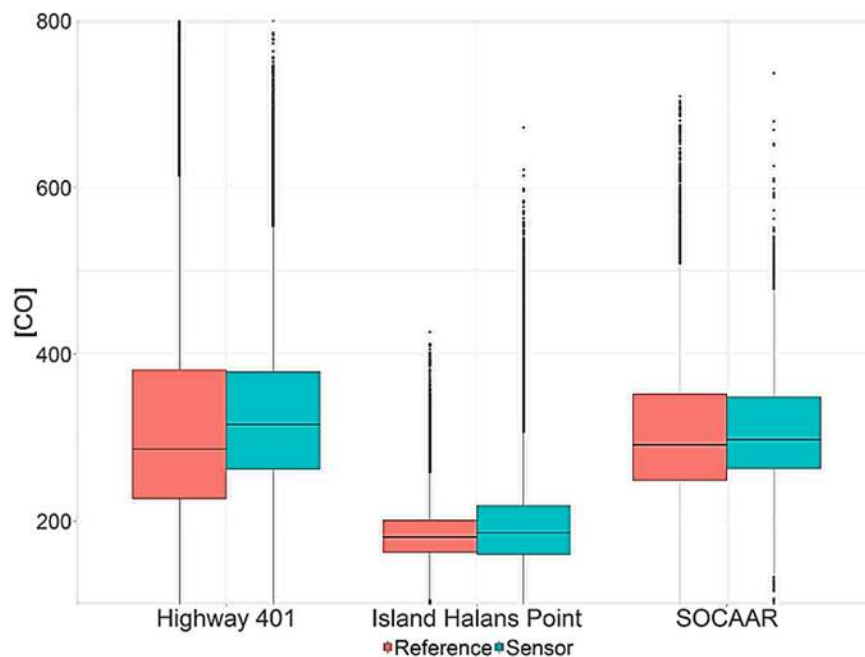


Figure 54. Assessment of the performance of CO measurement at SOCAAR, Highway 401 and Island sites. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-

minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

The NO_x levels were underestimated at sites with high level of traffic pollution. While earlier analysis indicated that calibration of NO_x models using data from high traffic locations could improve their performance, this is not always possible. Another method to improve model performance at higher pollutant levels is model predictor transformation.

To evaluate the extent to which predictor transformation can be used to improve the accuracy of MLR models in assessing the differences in pollutant distributions at different sites, MLR models were developed using transformed predictors. The most impactful transformed variables were selected based on the relative importance ranking output by `relaimpo` R with adjusted R squared used as a metric. The logarithm, exponential and square root transforms were ranked as the most important for NO_x. The impact of variable transformation on the accuracy of spatial analysis was evaluated by comparing the NO_x concentration distributions for reference monitors, original sensor device models and transformed sensor device models.

The ability of AirSENCE devices to correctly classify sites based on the level of NO_x was improved when predictor variables were transformed (Figure 55). Without variable transformation, the local SOCAAR and island NO_x levels are often higher than those at the highway site. Variable transformation helps develop regression models that correct for sensor nonlinearity and lead to more accurate estimation of inter-quartile range. For example, the interquartile range was smallest for Toronto Island site, medium for SOCAAR and high for Highway 401 for both reference monitors and MLR models derived using transformed predictors but not for the MLR models derived from raw predictor measurements. Despite the improved ability in ranking sites, the interquartile range is still larger at sites with low and medium traffic and smaller at sites with high levels of traffic pollution compared to NO_x reference interquartile range. PM_{2.5} models developed using transformed predictors also yielded significant improvement in the ability of the AirSENCE devices to correctly rank sites based on the level of pollution.

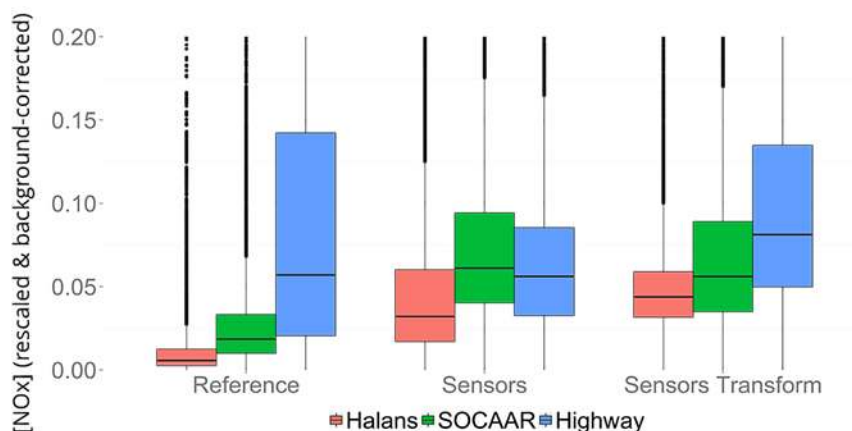


Figure 55. Comparison of distributions of NOx at the island, SOCCAR and highway sites for regression models derived using transformed and raw data. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.6.3 Impact of time on spatial variation

Multiple factors influence the similarity and differences in pollution distributions at sites, including type of sources present, time of day, day of the week, season, temperature, pollutant concentration and the levels of interferences. Binning sensor data can help evaluate sensors with respect to their ability to resolve the differences between sites. For example, knowledge of the conditions under which the sites are most divergent can help evaluate the ability of the AirSENCE devices to reveal spatial gradients.

Alternatively, the conditions when sites are known to be more similar to each other can be used for correction of bias and drift. Temporal variation in pollutant levels over a day or week is generally consistent across sites and is a promising factor for establishing the periods when the AirSENCE devices are more similar or different from each other. Therefore, one of the major goals of the analysis was to investigate how time of day affects the differences between sites.

The impact of time of day on the discrepancy between sites was analyzed by comparing the NOx distributions at SOCAAR to 5 different sites spread across Toronto during SCULPT campaign. MLR models were used to convert sensor data to NOx concentration. The sites varied in the urban density, proximity to the road, size of local roads as well as the average volume of local traffic. For the purpose of this

analysis, morning was defined as period of time between 5 AM and 11 AM, afternoon was defined as 11 AM – 5 PM, evening was defined as 5 PM – 11 PM and night was defined as 11 PM – 5 AM.

Nonparametric Kolmogorov-Smirnov Two Sample Test was used to quantify the difference between site distributions and assess the significance of the difference between sites. This test was selected because it makes no assumptions about the probability distributions of the predictors and bidirectional test was used because it makes no claims about the direction of the difference between sites. The difference between sites was measured by Kolmogorov-Smirnov D statistic, which increases as the difference between pollutant distribution at the compared sites increases.

It was found that the between site differences in NO_x distributions are greater in the morning compared to afternoon (Figure 56). The difference from the central site is statistically significant (Kolmogorov-Smirnov p value < 0.05) for all site comparisons but is especially pronounced for sites 10 and 15. This suggests that comparing pollutant levels during morning rush hour might be the best way to reveal spatial variation due to traffic-related air pollution (TRAP).

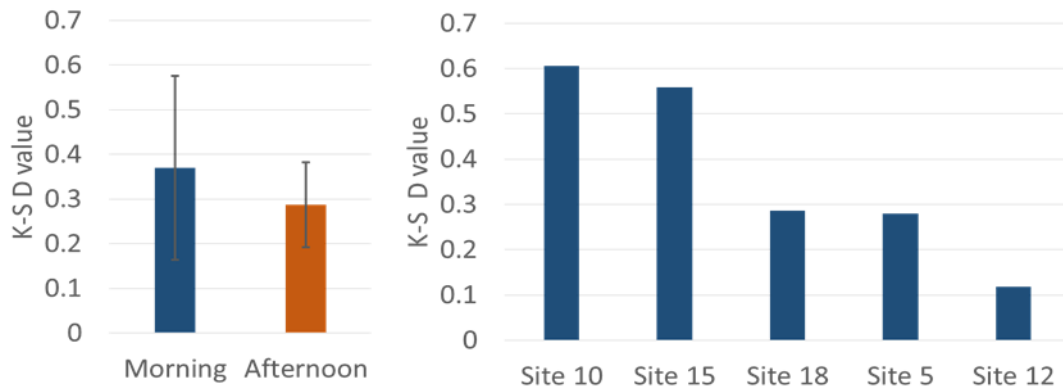


Figure 56. Analysis of the impact of time of day on divergence between central site and other sites.

Summary of the similarity between central site and other sites in the morning is also shown.

Kolmogorov-Smirnov D statistic is used as a metric of divergence. Morning is defined as 5 AM-11 AM.

Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

To verify the impact of time of day on the variation in traffic related pollutant distributions, distributions of NO_x at the sites with low medium and high level of traffic were compared. The devices were

deployed at SOCAAR, Toronto Island and Highway 401 MOE stations between December 2015 and February 2016. The sensor device NO_x distributions were compared for morning and afternoon as well as weekend and weekday.

It was found that differences across sites are more significant in the morning and on weekdays compared to those in afternoon and on weekends (Figure 57). Median, 25th and 75th percentiles were lower in afternoon and on weekends for high and medium traffic sites. The pollutant levels at low traffic site (Toronto Island), on the other hand, increased slightly in the afternoon and on weekends suggesting that increase in pollution at that site is more likely due to city background levels rather than local pollution sources. The higher levels of traffic during morning rush hour and on weekdays can explain the higher differences between sites.

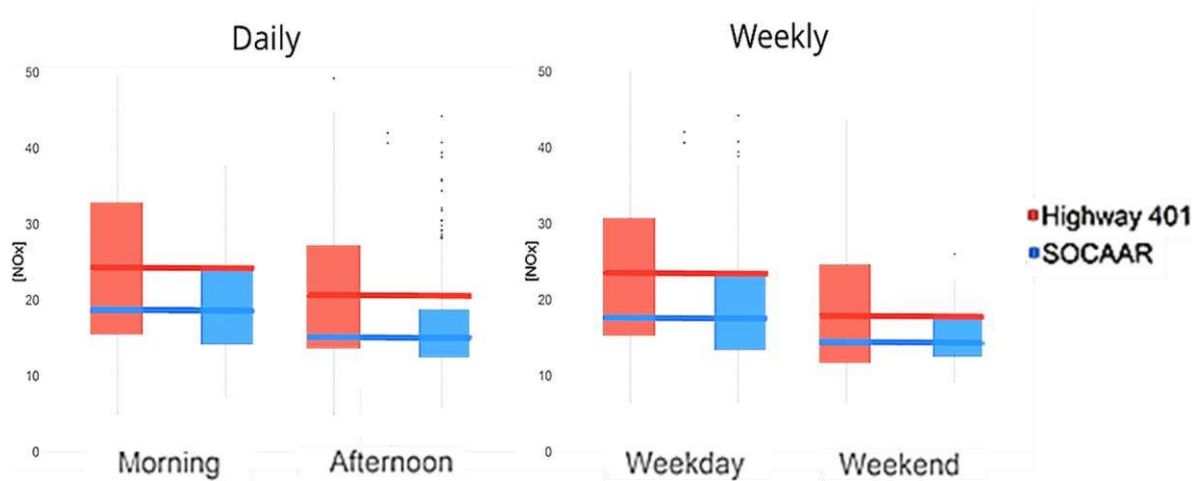


Figure 57. Analysis of the variation in divergence between sites depending on time of day and day of the week. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.6.4 Impact of wind on spatial variation

The differences between sites have been shown to vary depending on time of day and day of the week. Another factor that is known to have large impact on variation between sites is wind direction. Change in wind direction can change the influence of local sources on pollution levels at a site and therefore can highlight the differences between sites. Therefore, wind measurements not only help resolve differences between sites but can also highlight influences of specific sources.

The impact of wind direction on traffic-related pollutant distribution was investigated by assigning a wind direction category to each sensor device observation. Pollution observations associated with wind from 45-315-degree direction was categorized as coming from north, 45-135-degree direction – coming from east, 135-225-degree direction – coming from south and 315-225-degree direction – coming from west. The analysis showed that east, west and south were the predominant wind directions for the investigated time period and therefore only the observations associated with wind coming from these directions were considered. The spatial variation in both NO_x and CO was considered.

The analysis showed that wind from west is associated with highest pollutant levels and largest differences between sites for both NO_x and CO (Figure 58). This finding is consistent with these sites being downwind of more traffic emissions when winds are from the west and northwest. The spatial variation in NO_x was more impacted by wind direction than that of CO. Wind from east is associated with highest similarity between sites for NO_x and wind from south is associated with highest similarity between sites for CO. Binning the sensor device measurements based on wind direction can help improve the sensitivity of the analysis of spatial variation of air pollution.

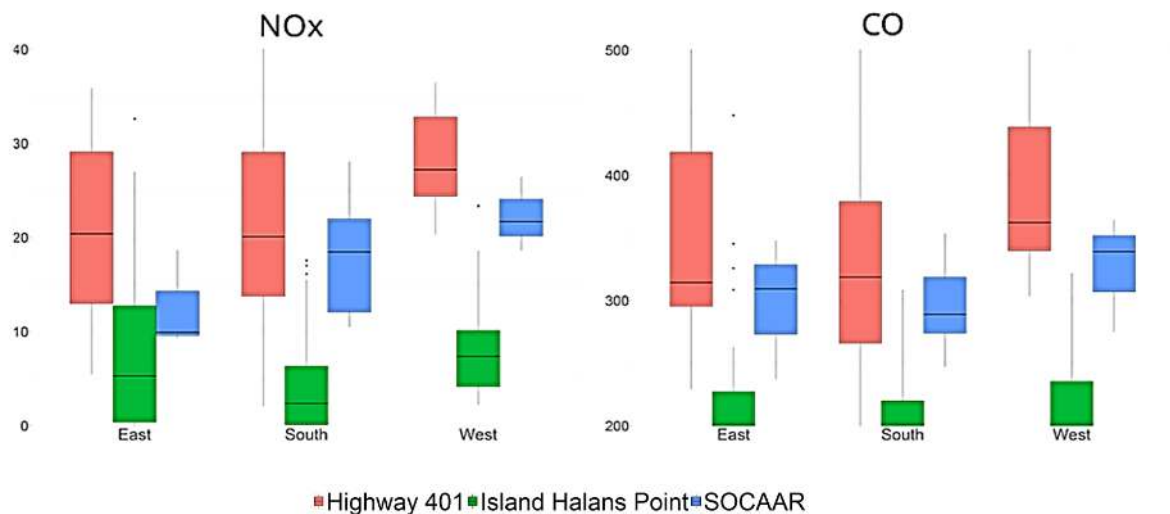


Figure 58. Analysis of the variation in divergence between sites for NO_x and CO depending on wind direction. Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

4.6.5 Summary

In summary, it was found that spatial variability in pollutant levels can be resolved and the difference in magnitude of pollutant levels measured by AirSENCE devices is supported by the literature LUR models. AirSENCE devices were able to resolve absolute differences between sites both in terms of average pollutant levels and the peak levels.

Sensor device models for CO model showed the best performance in detecting differences between sites compared to other pollutants. NO_x and CO₂ estimates by the AirSENCE devices could only be resolved at medium and low traffic sites, while O₃ estimates by the AirSENCE devices could only be used to compare medium and high traffic sites.

The ability to resolve sites and site ranking also varied over the course of day and week. The differences across sites were more significant in the morning and on weekday compared to those in afternoon and on weekend. Therefore, comparing pollutant levels during week day morning rush hour might be the best way to reveal the spatial variation of traffic-related pollution.

Wind from west was found to be associated with highest pollutant levels and largest differences between sites for both NO_x and CO. Binning the sensor device measurements based on wind direction can help improve the sensitivity of the analysis of spatial variation of air pollution.

The ability of AirSENCE devices to correctly classify sites based on the level of NO_x was improved when predictor variables were transformed, since it allowed for correction for sensor nonlinearity and lead to more accurate estimation of inter-quartile range. Overall, the AirSENCE devices provided pollutant patterns that were consistent with the known temporal and spatial patterns of traffic related air pollutants.

4.7 Traffic-related pollution

Source contribution analysis is one of the most promising applications of AirSENCE devices. Due to low cost, dense networks of AirSENCE devices can be deployed, and hold promise in improving understanding of types of local sources, the magnitude of their contribution as well as temporal variation in their impact. In addition, since source analysis is mostly categorical, the limitations of

regression models can be avoided. Source contribution analysis can even be conducted using the sensor data directly without the need to develop calibration models. Even when the regression model accuracy is insufficient for analysis of pollutant levels, the sensor device performance might still be acceptable for detecting differences in source contribution. A key goal of source analysis was to evaluate ability of AirSENCE devices to rank sites based on level of traffic and identify the analysis methodology and metrics which can improve the ability of AirSENCE devices to resolve the contribution of local traffic sources. Another major goal was to understand the extent to which the multipollutant patterns and combining sensor data with wind direction measurements can help improve source contribution assessment. Multipollutant pattern analysis can also help distinguish traffic source types like diesel and gasoline.

To assess the ability of AirSENCE devices to detect traffic-related pollution, temporal patterns were compared at sites with different levels of traffic. The impact of methods such as background subtraction on ability to detect traffic pollution peaks was investigated. The relationship between traffic volume and traffic related pollution levels was tested. Sources contributing to pollution levels at the site were identified by comparing the correlations between the pollutants at the sites. By combining wind direction data with sensor device measurements, source contribution was evaluated, and major local sources were ranked based on relative contribution. The variation in local source influence over the course of a week and 5 months was evaluated to understand the frequency and magnitude of source impact variation. Relative contribution of local sources was compared for major traffic pollutants and sites were ranked based on the extent of local source contribution. Sensors were also tested on their ability to distinguish different types of traffic pollution (i.e. gasoline and diesel) sources by quantifying the ratio of pollutants associated with these two sources. Change in co-occurrence of peaks over the course of a week was used to identify trends in source contribution. In addition, the comparison of average daily and weekly trends in response of NO_x, VOC, CO and PM sensors was used to validate the temporal patterns in contribution of sources and assess the potential applications on uncalibrated sensors to distinguish source patterns.

The following questions were addressed:

Spatial variation in traffic pollution

1. Can sensor device pollutant estimates be used to distinguish between sites based on their relative levels of traffic pollution?

2. Is the average pollutant concentration related to average traffic volume within a 1500m buffer of a site?

Multi-pollutant patterns and source identification

3. Do correlations between different sensor device pollutant models improve the ability to compare pollutant sources?

Impact of wind direction on source contribution

4. Can wind direction data improve the ability of AirSENCE devices to identify local pollution sources?
5. How does the pollution source contribution vary over the course of a day and season?

Local and background source contributions

6. How do pollutant models compare with respect to local and background source contribution?
7. How do the site rankings based on local pollutant source contribution vary for different types of pollutants?

Traffic source categorization

8. Can AirSENCE devices be used to compare sites based on type of traffic source contribution?
Can AirSENCE devices be used to resolve gasoline contribution from diesel contribution?
9. Can the temporal variation in diesel and gasoline pollution sources be detected by AirSENCE devices? Is it possible to resolve source contribution using raw sensor output directly without calibration?

4.7.1 *Local vs background impact*

Air pollution monitoring campaigns have indicated that air pollutant levels vary across sites and show dependence on traffic volume and fleet characteristics. The number, location and size of the nearby roads affects the air pollution level as well as timing and frequency of increase in air pollution concentration. However, the use of road metrics (i.e. road length and width) is often insufficient for evaluating the impact of traffic on pollution levels at a site. Many other factors impact the variation in air pollution, including wind direction, wind speed, mixing height and traffic conditions. These factors can be highly variable and inconsistent between days and contribute to large variation in pollutant levels as well as variability in relative contribution of local sources. For example, during Pan Am Game campaign the pollutant levels at Old Toronto site were often high due to proximity to Gardiner Expressway, but in the morning the levels were often much lower compared to downtown sites like

SOCAAR. In addition, the short-term variation in pollutant levels and the frequency of change was often higher at the downtown site suggesting large variability in local source contribution (Figure 59).

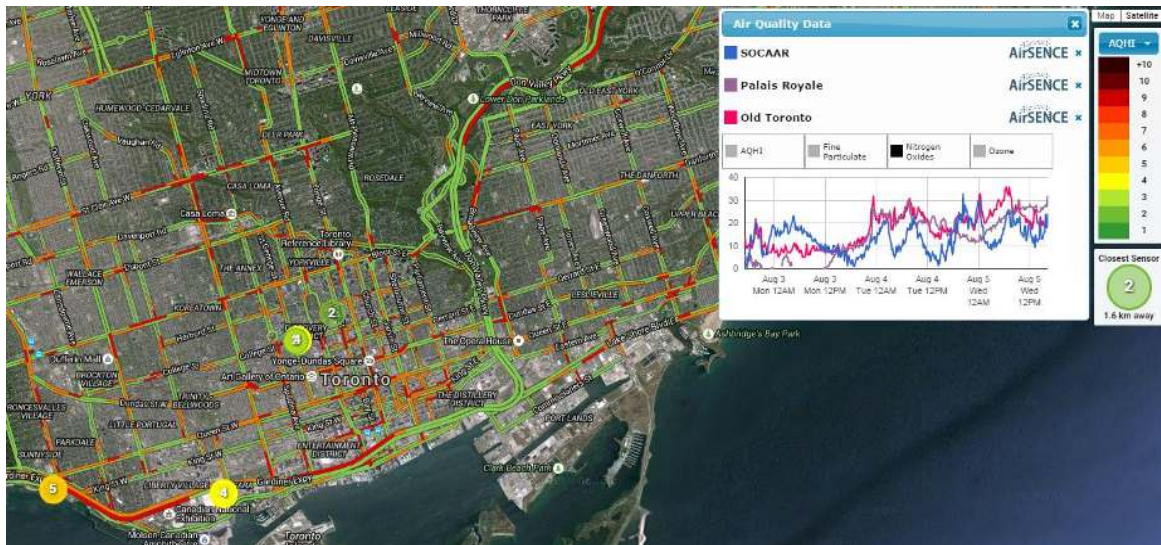


Figure 59. Pollutant levels estimated by AirSENCE devices vary over the course of the day and depend on the amount of traffic. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.

When temporal patterns are closely assessed, two components of temporal variation emerge: a background component and a local component (Figure 60). The temporal variation in background pollutant levels occurs at a time scale of several hours and is the result of meteorology and contribution of city-wide and regional sources. The temporal variation in local pollutant levels occurs at a time scale of seconds-minutes and is contributed by local sources. It is expected that when the sites are compared at high temporal resolution (i.e. 1min or lower), local patterns would emerge helping characterizing sites and better evaluate differences between them.

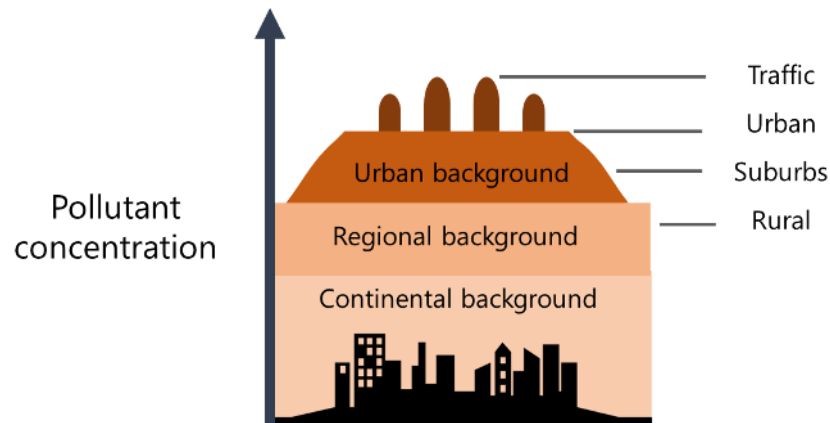


Figure 60. Pollution component contributions (inspired by Fuzzi et al. 2015).

The contribution of traffic to pollutant levels at a site was evaluated by comparing the temporal variation in NO_x at 5 different sites spread across Toronto during the SCULPT campaign. The sites varied in the urban density, proximity to the road, size of local roads as well as the average volume of local traffic. In order to analyze the local component of temporal variation (i.e. traffic pollution contributed by a nearby road), all the data was averaged to 1 min temporal resolution to retain high frequency pollution variation data while reducing the amount of noise. Background was quantified as 1 hour rolling minimum. This background signal was subtracted from the 1 min resolution data to calculate the local contribution to pollutant levels at a site. Sites with different levels of traffic were then compared to each other. Traffic volume was quantified in vehicles per day within a 1500m buffer. The spatial variation of both background and local source component of pollutant measurements was compared.

It was found that background subtraction helped emphasize the differences and similarities between sites in terms of local traffic source contribution (Figure 61). While both site 15 and site 5 showed large amounts of background variation throughout the day, the extent to which this variation was contributed by local traffic sources was unclear. Site 5 is a low traffic site that nonetheless has high NO_x background level. When the background is subtracted, the high-frequency variation in NO_x levels at site 15 is clearly greater than that at site 5. The variation in local source contribution at the sites becomes emphasized, especially during afternoon rush hour. Local component of pollutant distribution is therefore a more representative measure of the local source impact and better reflects local traffic.

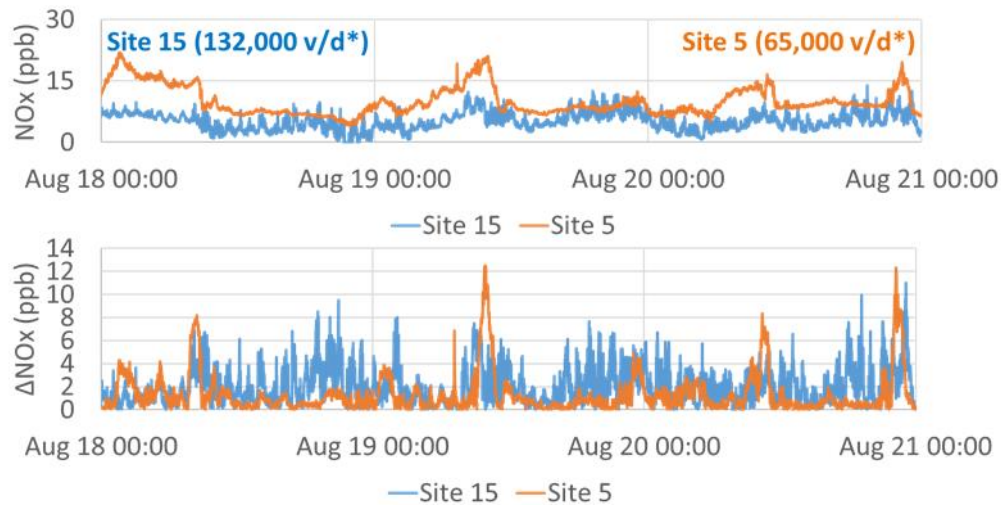


Figure 61. Analysis of local differences between site 15 (site with higher traffic influence) and site 5 (site with lower traffic influence) sensor measurements (1 min average data). Original sensor measurements are compared to those with background removed. *Traffic data is shown in vehicles per day within a 1500m buffer. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

To evaluate the extent to which local source contribution can be used to rank sites with different levels of traffic pollution, % local contribution was calculated for each of the 5 sites investigated during the SCULPT campaign. The normalized % local contribution was found by calculating relative proportion contributed by local sources and normalizing it by dividing the site measurements by those at SOCAAR. The normalized % local contribution was then compared to the average volume of traffic within a 1500m buffer of the site.

It was found that relative to the SOCAAR site the percent of local NO_x levels was highest at site 10 and site 15, medium at site 12 and 18, and lowest at site 5 (Figure 62). These trends were found to be consistent with local traffic level. High local contribution (>1% normalized local contribution) at SOCAAR, site 10 and 15 was associated with high traffic level (>100,000 vehicles/day), medium local contribution (0.75-1% normalized local contribution) at site 12 and 18 was associated with medium traffic level (~70,000 vehicles/day) while low local contribution (<0.75% normalized local contribution) at site 5 was associated with lower traffic level (~65,000 vehicles/day).

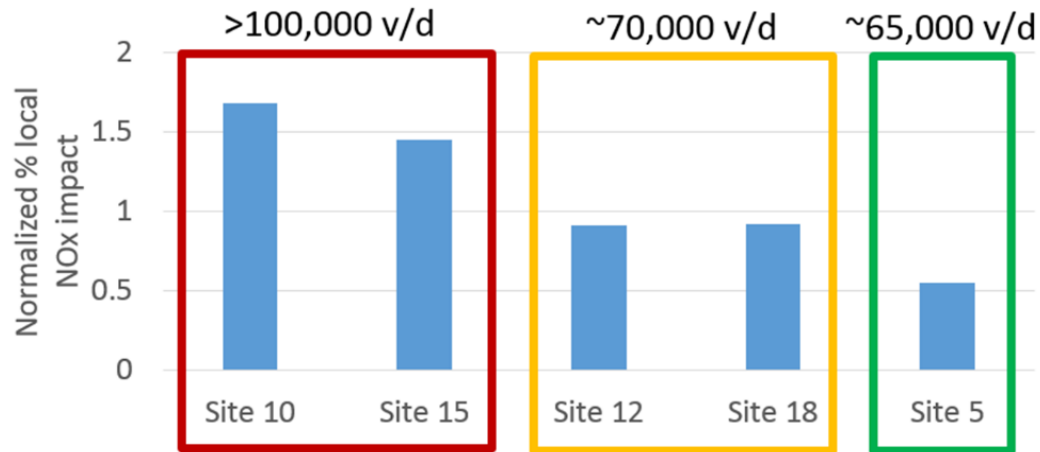


Figure 62. Comparison between local pollution source contribution and the traffic volume at different sites. *Traffic data is shown in vehicles per day within a 1500m buffer. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

4.7.2 Major source identification

As shown earlier, the multi-pollutant output of sensor array devices can be used to identify differences in pollutant levels and their changes in different regions of the city. This study also provided some evidence that multi-pollutant temporal patterns can be used to reveal differences in source composition. Ideally, analysis of local pollution mixtures through correlation of pollutant measurements at specific sites may be able to identify common sources and help improve understanding of the differences in sources between the sites.

Correlation analysis was conducted between NOx and other pollutants (i.e. CO, CO₂, O₃ and PM_{2.5}) at 3 different sites in Toronto during the SCULPT campaign. The analysis was conducted to test the hypothesis that differences in types of sources can be identified by examining multi-pollutant patterns. The degree of cross-correlation was quantified at each site and correlation pattern was used to evaluate whether the pollutants originated from a common source at that site.

Correlation between pollutant estimates was found to differ between sites and suggests different sources (Figure 63). As expected, traffic was a major contributor to pollution levels at site 18 and SOCAAR site given the higher correlation between NOx, CO and CO₂. On the other hand, sites with less

traffic (site 5) showed less correlation between traffic-related pollutants. PM_{2.5} is correlated with NO_x only at site 18, which might possibly be due to construction at that site during the period when measurements were collected. Overall, analysis of correlation between pollutants is indicative of dominant sources and is useful for site comparison.

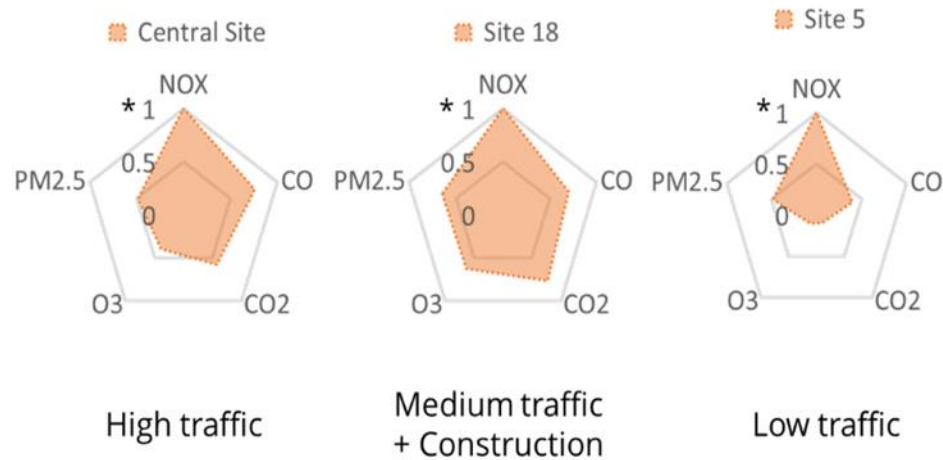


Figure 63. Comparison of pollutant signatures associated with different sites. Axes represent absolute Pearson correlation coefficients between NO_x and other pollutants estimated by AirSENCE devices at different sites (central site, site 18, and site 5). A larger shaded area indicates stronger correlation. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

Combining sensor device output with wind data can help understand the timing and magnitude of impact of sources at a site (Appendix 5: Wind impact analysis). A 3-dimensional analysis can be achieved by tracking the change in pollutant levels in response to changes in wind direction and over time while verifying whether the change is consistent across all pollutants. Wind direction measurements help identify potential local source categories, pollutant co-presence helps validate those categories and change over time helps identify the extent to which they are consistent.

Analysis of sensor response as a function of wind direction was conducted by binning the sensor data relative to wind direction. Wind speed and direction data from the meteorological station on the roof of the Wallberg building was used. Only the measurements corresponding to wind speed more than 0.2 and less than 3 meter per second were selected, since the very low and high wind speed measurements

are known to be associated with higher uncertainty. The wind direction data was binned into 12 sectors and the pollutant data was binned in 6 categories. For each of the 12 wind direction sectors, the average contribution of each of the 6 pollution level categories was then calculated and the change over time was examined.

Traffic pollution is known to be highly variable over the course of a day. While the amount of traffic itself varies as a result of rush hour and daily commute patterns, wind direction is another important factor affecting the level and type of pollutants measured at a site. The combined impact of time of day and wind direction on pollutant level was assessed by binning the pollutant data into 4 categories: morning, afternoon, evening and night. For the purpose of this analysis, morning was defined as period of time between 5 AM and 11 AM, afternoon was defined as 11 AM – 5 PM, evening was defined as 5 PM – 11 PM and night was defined as 11 PM – 5 AM.

Increase in levels of NO_x in the morning was found to be associated with wind from the south and south-west, increased NO_x in afternoon was associated with wind from south and east, increased NO_x in the evening was associated with wind from south and north-west, while higher level of NO_x at night was associated primarily with wind from north (Figure 64). This finding suggests that the background and local sources of NO_x are different. The results are also supported by the typical traffic patterns in Toronto. For example, the increased pollution levels associated with north-west wind might be related to increased traffic at the west end of Gardiner Expressway in the morning (Figure 65). Increase in traffic along Yonge street as well as downtown portion of Gardiner Expressway in the afternoon might be the cause of increase in pollution levels associated with east and south winds at that time of day. Larger contribution associated with wind from the north in the evening and at night might be due to increased freight, construction and city maintenance diesel vehicle volume, which is often scheduled in the evening and at night to avoid disruption to traffic.

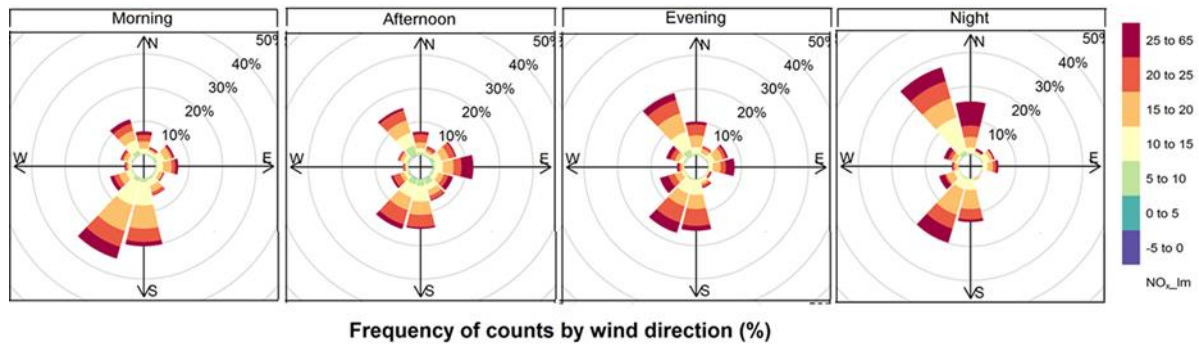


Figure 64. Variation in pollution source contribution over a course of a day. The graph is split into panels to show the dominant wind direction and speed for different times of day: morning, afternoon, evening and night. The NO_x_lm colour scale indicates the NO_x concentration measurements by the sensor device. Wind speeds are split into the intervals shown by the scale in each panel. The grey circles show the % frequency occurrence based on 1 min data. The plots show the proportion (here represented as a percentage) of time that the wind is from a certain angle and wind speed range. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.

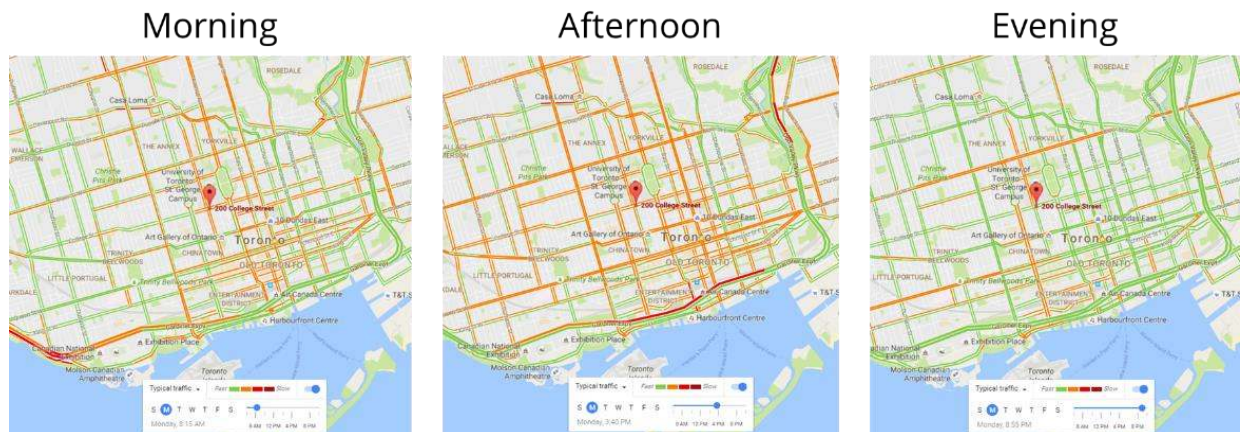


Figure 65. Variation in traffic volume at SOCAAR in the morning, afternoon and evening.

Wind direction also affects long-term variation in pollution levels. Whether influenced by changes in local sources (i.e. nearby construction), changes in mixing height or pollutant transport distance, the wind direction associated with increase in traffic-related pollution changes over time. To evaluate extent of its change, sensor data was binned by month between May and September 2015. Change in wind direction associated with the increase in NO_x level was examined.

The results indicated that the wind direction associated with increase of NO_x levels changed over the course of summer (Figure 66). While south-west and east wind dominated the increase in NO_x levels in May and June, north-west sources had a much bigger effect in July and August. The impact of south wind increased in August and September, while the north-west wind continued to have a large impact. These results suggest that the influence of local sources on pollution levels at a site varies over the course of a season and the monthly variation in source impact is big enough to be resolved by AirSENCE devices. The lack in consistency of source impact with time also suggests that wind direction data is important for effectively comparing sites based on the influence of local sources.

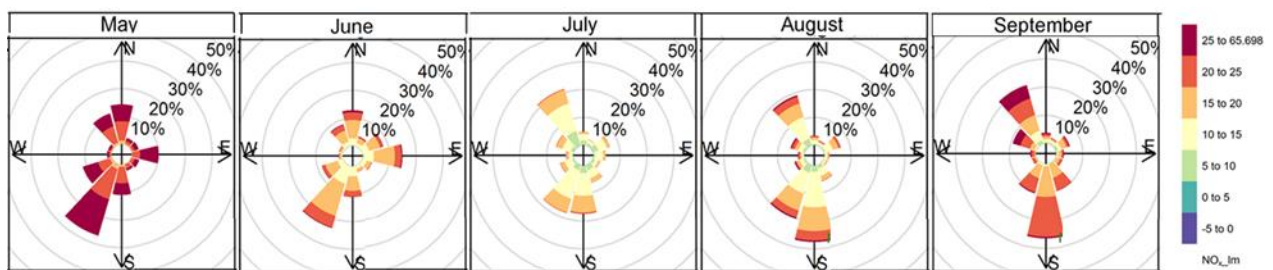


Figure 66. Variation in pollution source contribution over a course of 5 months. The graph is split into panels to show the dominant wind direction and speed for different months. The NO_x_Im colour scale indicates the NO_x concentration measurements by the sensor device. Wind speeds are split into the intervals shown by the scale in each panel. The grey circles show the % frequency occurrence based on 1 min data. The plots show the proportion (here represented as a percentage) of time that the wind is from a certain angle and wind speed range. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.

4.7.3 Traffic source comparison

Motor vehicle traffic (personal and freight) is a major contributor to Toronto air pollution and the associated health effects. Major components of vehicle exhaust include CO₂, NO_x, CO, hydrocarbons, black carbon (BC) and PM. While gasoline car emissions contain higher proportion of CO and hydrocarbons, diesel truck emissions contain more NO_x, BC and PM (Figure 67). Despite higher levels of NO_x and PM in diesel exhaust, they're also present at significant levels in some gasoline emissions. CO₂ is present at high concentration in all vehicle exhaust and is often used to understand the total amount of emissions and derive emission factors.



Gasoline	Diesel
Higher CO Higher Hydrocarbons	Higher PM Higher NOx
	

Figure 67. Differences in traffic emissions associated with gasoline and diesel vehicles.

Differences in the pollutant composition of diesel and gasoline sources suggests that the comparison of output of regression models for different pollutants might help resolve sites relative to the contribution of the different types of vehicles. For example, comparison on local contribution of NO_x and CO and CO₂ at different sites may both help quantify relative amounts of diesel traffic sources and suggest the magnitude of the impact of local sources relative to background sources. For example, higher local contribution of NO_x compared to CO might indicate more diesel sources compared to gasoline sources, while high frequency pollutant variation might suggest large relative impact of nearby roads.

The % local contribution was calculated for NO_x, CO and CO₂ for each of the 5 sites investigated during the SCULPT campaign. The pollutants were then compared based on the average local contribution across all sites. In addition, the normalized % local contribution was found for each site by calculating relative proportion contributed by local sources and normalizing it by dividing the site measurements by those at SOCAAR. Sites were then ranked based on the influence of local and background sources for each pollutant.

It was found that NO_x showed highest local contribution across sites (24%), followed by CO (7%) and CO₂ (2%) (Figure 68). NO_x is less stable and has a much shorter atmospheric lifetime than CO and CO₂, which as a result are transported over longer distances and have larger background levels. The relative background and local contribution of NO_x, CO and CO₂ was also found to vary between sites. Sites 10 and 15 are more influenced by local sources compared to SOCAAR, sites 12, 18 and 5. Larger local contribution of NO_x relative to CO at sites 10 and 15 suggests that these sites are more impacted by

diesel vs gasoline pollution than the other sites. The local diesel source contribution relative to CO at site 10 is highest compared to other sites, since the ratio between normalized local contribution on NOx and CO is the above 3 for that site and below 1 for all other sites. Conversely, at site 12 higher local CO contribution relative to NOx suggests a larger influence of gasoline vehicle sources. An interesting feature of background sites is higher relative local contribution of CO₂ compared to NOx and CO, suggesting generally cleaner vehicles with lower emissions of NOx and CO per mass of fuel burnt. These results suggest that AirSENCE devices are sensitive and selective enough to allow for ranking of sites based on the local contribution for different pollutants. Extracting local pollutant contribution can help cluster sites based on the level and types of pollution sources.

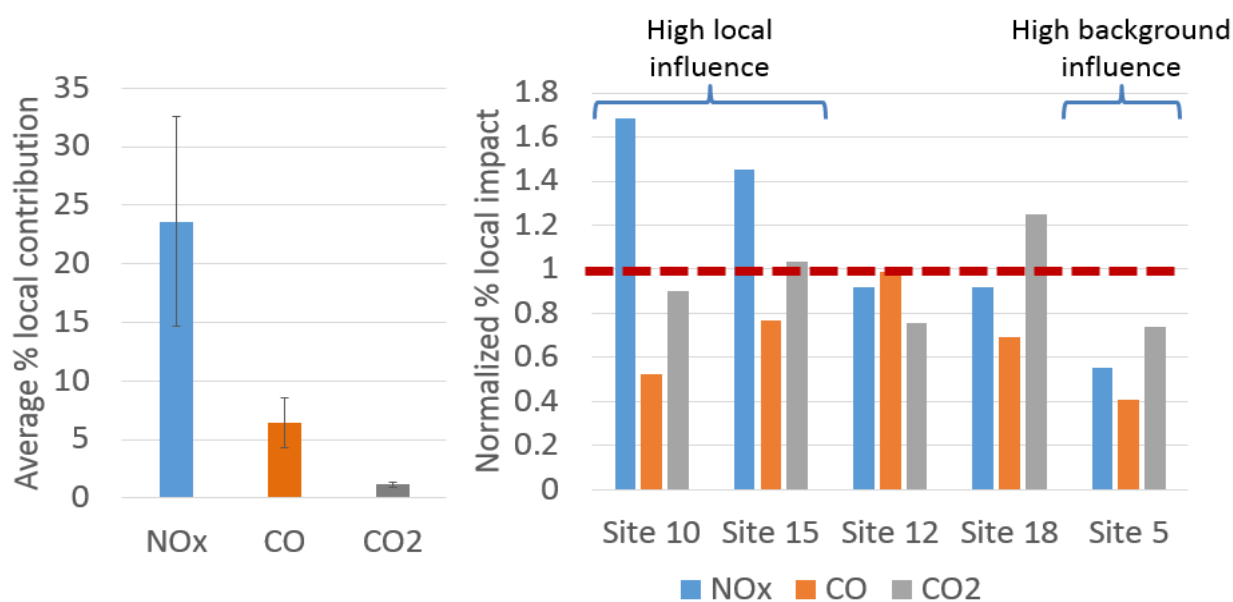


Figure 68. Assessment in variation of local source contribution for 3 traffic-related pollutants: NOx, CO and CO₂. Ranking of sites based on local traffic influence and comparison of 3 traffic-related pollutants is also shown. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

Analysis of local component of pollutant estimates by AirSENCE devices has been shown to be useful for qualitative site comparison. However, deriving quantitative cumulative comparison metrics from multi-pollutant data can be more useful for real-time analysis. For example, traffic-related pollutant ratios can simplify the comparison of source composition at different sites. In addition, pollutant ratios between

sites can suggest the extent to which sites are similar to each other. In particular, emission factors can be derived to help compare sites with different source composition.

Analysis of pollution peak ratios and emission factors requires the comparison of plume peaks for the key traffic-related pollutants like NO_x, CO and CO₂. Therefore, the similarity of peak patterns was first compared for AirSENCE devices for these three pollutants. Background subtraction was found to be essential for comparison of plume peaks for NO_x, CO and CO₂. At site 18, most peaks coincided very well in background-corrected data, suggesting promise in using sensors to estimate emission factors (Figure 69). While NO_x, CO & CO₂ have similar high time frequency patterns, the relative background contribution differs for each pollutant. These differences in background contribution obscure the relative local source contribution. Therefore, it was essential to perform background subtraction.

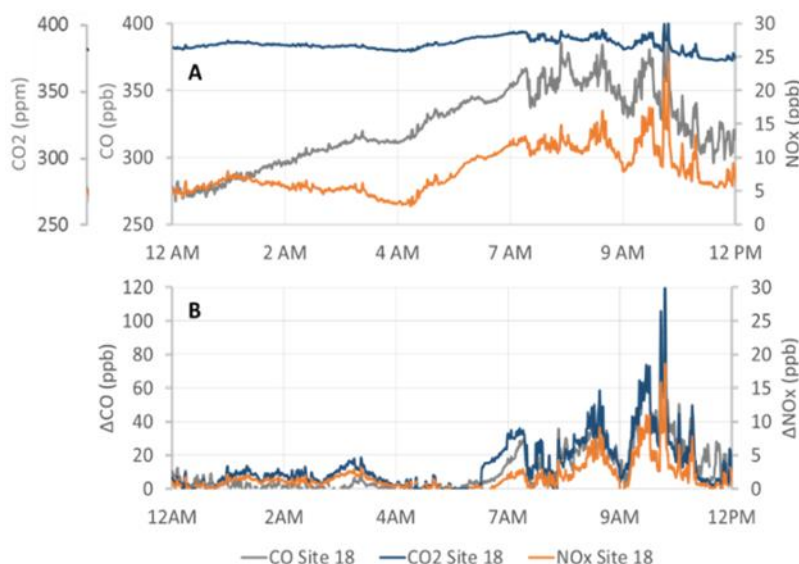


Figure 69. Comparison of diurnal trends for CO₂, CO and NO_x at site 18 for original data (A) and background-corrected data (B) (1 min average data. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

Since CO and NO_x are both major markers of vehicle exhaust, it was hypothesized that the co-occurrence of NO_x and CO peaks could indicate periods of the day when peaks in these pollutants are caused by traffic (Appendix 6: CO/NO_x ratio analysis). Comparison of the CO and NO_x peak co-occurrence over time can help understand traffic pollution patterns over the course of a day and a

week. Site 18 was characterized by the broadest distribution in CO/NO_x ratios and therefore was the focus of detailed time pattern analysis.

Background-subtracted data was used for analyzing the similarity in temporal patterns of NO_x and CO at site 18. The comparison was conducted over 4 days – 2 weekdays and 2 weekends. Both the frequency and the amplitude of peaks was investigated. In addition, the average CO/NO_x ratios were also compared at site 18 on weekdays and weekends. CO/NO_x ratios were then calculated for a set of representative peaks, which were coincident for the two pollutants. Background-subtracted data was used for calculating the ratios. Peaks were identified by calculating a smoothed first derivative and identifying when it crosses zero. In order to ensure that the selected peaks correspond to actual vehicle plumes, amplitude and slope thresholds were also applied to filter out the peaks that were too small or too wide. The peaks which were coincident for CO and NO_x were then selected and CO/NO_x ratios were found for each peak.

The CO/NO_x ratios were found to vary depending on the day of the week (Figure 70). On weekdays, the NO_x and CO peaks were very similar in frequency and amplitude, particularly in the morning and evening. Morning period was characterized by highest increase in both pollutants, validating the importance of morning period for understanding spatial variability in source contribution. However, on weekends there was no corresponding increase in NO_x levels when the CO levels increased in afternoon. The weekend increase in CO/NO_x ratio is likely the major cause for the broad distribution in CO/NO_x ratios at site 18. A possible cause for the observed pattern could be the relatively high weekend gasoline traffic compared to other sites, but reduced diesel traffic on weekends compared to weekdays. The observed increase in CO/NO_x ratio on weekend was larger compared to that detected at SOCAAR site (where the CO/NO_x ratio is ~1.18 times higher on weekends (Sofowote et al. 2017)). These results provide additional evidence to support the site-specific nature of temporal variation in local source composition.

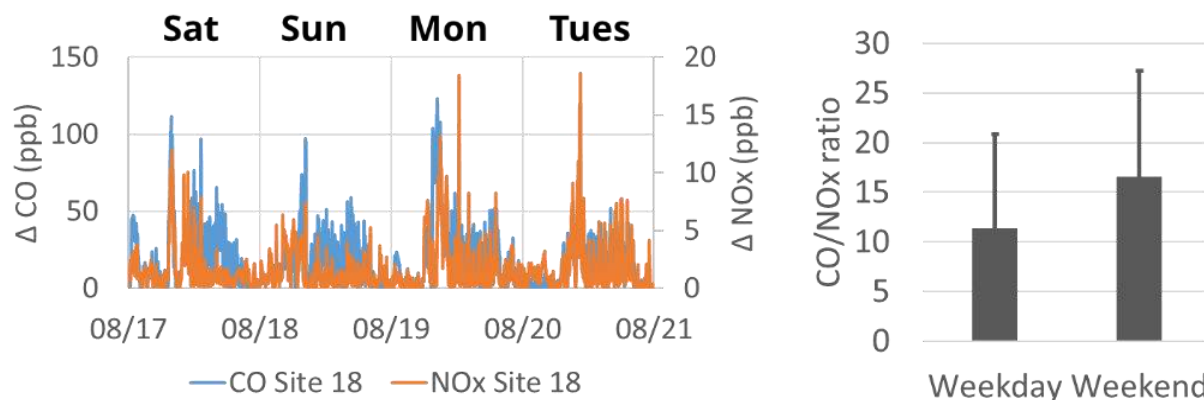


Figure 70. Time series of local contribution to NOx and CO levels at site 18 (1 min average data). Comparison of relative levels of CO and NOx for weekday and weekend are shown. Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

Analysis of traffic contribution to pollution levels at a site is generally a two-step process. Initially, the pollutant estimates are derived from sensor readings by finding and applying optimum calibration models. In the second step, the signal relevant to traffic is extracted through background subtraction and the relationship is found between pollutant regression model output and traffic metrics (i.e. emission factor, pollutant ratio, site ratio). The output of this process is confounded by the propagation of errors in both steps of the analysis. Earlier analysis has indicated that estimation of exposure metrics directly from sensor data can help improve the accuracy and reproducibility of the result. For example, the method for deriving the Air Quality Health Index (AQHI) directly from sensor data outperformed the alternative method of calculating the pollutant concentration and applying the AQHI formula. Traffic source analysis using the sensor data directly also has a potential to reduce the analysis error and improve accuracy and reproducibility. In addition, this method can improve the quality of the analysis by making use of the sensor data that is currently not used in regression models. For example, some volatile organic (VOC) sensors are informative for analysing sources but are not currently used in regression models because of limited VOC reference data.

The ability of sensors to detect meaningful temporal pollution patterns was tested by investigating average daily and weekly patterns at SOCAAR site (Device C6, site C), west site (Device C11, Site 1 at Palais Royale), east site (Device C10, Site 2 in Ajax) and north site (Device C1, Site 3 in North Toronto). The data was collected in May-September 2015 during a multi-site deployment during Pan Am Games in Toronto. Average temporal variation over the course of a day and a week was investigated for 4

categories of sensors: 1) NO_x and O₃ sensors, 2) CO sensors, 3) PM_{2.5} sensors and 4) VOC sensors. The magnitude of relative sensor response as well as consistency of that change across the different sensors was used to evaluate the timing of change in the impact of gasoline and diesel sources.

The average sensor response was found to differ for NO_x, CO, PM and VOC sensors over the course of a day and a week (Figure 71). Daily pattern in sensor response was found to be similar across sites for most sensors, while weekly patterns varied more between sites. These results suggest that analysis over the course of a week is more informative for comparing sites relative to background pollution levels. Higher time resolution analysis and background subtraction would also help resolve local source contribution. Several patterns emerge with respect to gasoline and diesel source comparison. Gasoline fuel is known to contain higher concentration of CO and hydrocarbons, while diesel fuels contains higher concentration of NO_x and particulate matter. Over the course of a day, largest increase in both CO and VOCs was found to occur in the morning between 6 AM and 12 PM, corresponding to the morning rush hour traffic. Trends in diesel-associated pollutants are not as consistent. While NO_x and O₃ sensor response is highest in the evening and at night, the particulate matter sensor response increases during the day. Both NO_x and particulate matter sensors indicate slight increase in response 12 – 3 AM suggesting that this increase could be contributed by increased logistics truck traffic at night. It was also found that diurnal patterns in CO and PM_{2.5} sensor response are similar, suggesting a common background source. In addition, NO_x sensor response increase after 4 PM co-occurred with VOC sensor response increase at that time suggesting that evening source composition might be different from that in the morning. Over the course of a week, all of the traffic related pollutants were found to be increase on Tuesdays and Sundays at most sites. Overall, raw sensor response was found to be sufficient for detecting differences in source composition at different times of day and over the course of a week. However, higher time resolution analysis is required for distinguishing pollutant patterns at different sites.

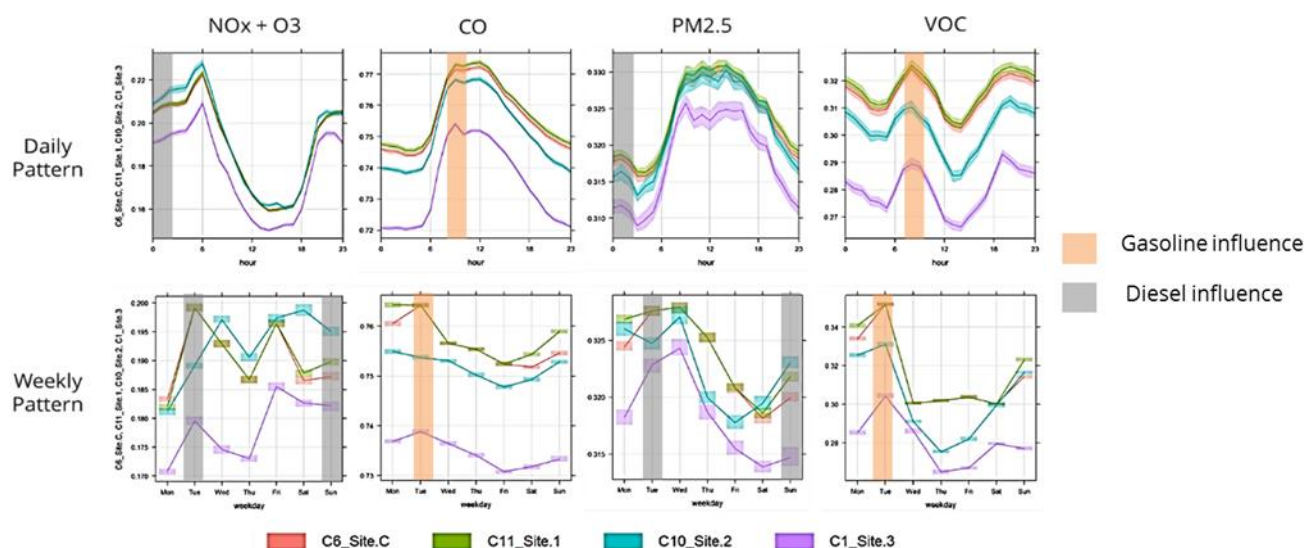


Figure 71. Diurnal and weekly trend clusters for 4 types of sensors (inorganic gas, carbon monoxide, organic gas and particulate matter) at 4 different sites are shown. Raw sensor voltage output was normalized at each site prior to comparison. Only relative changes displayed are meaningful, not absolute levels. Original data was collected between April 20, 2015 and September 9, 2015, sampled at 1-minute resolution, 191812 observations, 90% of data used. Data source: Pan Am Campaign, Summer 2015, AirSENCE Rev2.

4.7.4 Summary

Background subtraction was found to help emphasize the differences and similarities between sites in terms of local traffic source contribution. Local contribution is more representative of local traffic. For example, local NOx level was found to be related to local traffic level and could be used to distinguish sites with low, medium and high level of traffic.

Multi-pollutant patterns were found to be useful for detecting differences in sources. Comparison of the degree of correlation between different pollutants measured by AirSENCE devices revealed differences in source composition at different sites.

While NOx, CO & CO₂ have similar high time frequency patterns, the relative background contribution differs for each pollutant. NOx was found to have the highest local contribution across sites, followed by CO and CO₂. The devices were also found to be sensitive and selective enough to allow for ranking of sites based on the local contribution for different pollutants as well as the level and type of pollution

sources. Relative levels of NO_x and CO varied spatially and temporally suggesting that changes in vehicle fleet composition can be detected by AirSENCE devices.

The average sensor response was found to differ for NO_x, CO, PM and VOC sensors over the course of a day and a week. Daily pattern in sensor response was found to be similar across sites for most sensors, while weekly patterns varied more between sites. Temporal patterns suggest highest diesel contribution at night and highest gasoline contribution – in the morning. Overall, raw sensor response was found to be sufficient for detecting differences in source composition at different times of day and over the course of a week.

5 Conclusions

The studies presented in this thesis revealed several key insights about the performance of low-cost sensors, accuracy and precision of pollutant models, reproducibility of sensor performance over time and in response to interferences as well approaches for improving sensor accuracy and reliability. In addition, detailed analysis is presented on the opportunities to apply sensor array devices for analysis of temporal and spatial patterns, as well as assessment of air quality index and traffic-related pollution.

Analysis of a large range of gas sensors revealed several key challenges. One of the biggest issues is the high inter-device sensor variability in sensitivity caused by limited quality control of low cost commercial sensors. In addition, the variation between sensors increases with increase in temperature and concentrations of pollutants like CO and O₃. Another major issue is sensor response nonlinearity as indicated by left-skewed sensor distribution and reduced sensitivity at high pollutant levels. Sensors were also found to be cross-responsive, which is both a challenge and an opportunity. While sensor cross-responsiveness needs to be corrected for, it can be useful for detecting patterns associated with complex air pollution mixtures.

It was found that multiple regression models can provide more accurate pollutant estimates when multiple different sensor predictors are used compared to models with single sensor predictors. Sensor nonlinearity was one of the biggest contributors to reduction in pollutant model accuracy, which was found to be lower than measurement precision. While precision improved with increase in pollutant concentration, accuracy was highest for mid-range pollutant concentrations, partially because of the higher frequency of these pollutant levels in the calibration data. NO_x and PM_{2.5} models were found to have lower accuracy and precision compared to other pollutants. Air quality health index (AQHI) estimation from sensor readings was successfully demonstrated and accuracy was found to be higher with direct regression model fitting compared to calculation using AQHI formula and pollutant estimates. Sensor array devices also facilitated evaluation of variation of AQHI in space and time.

Three aspects of device reproducibility were evaluated: drift over time, impact of interferences and impact of site-specific mixtures. Drift was found to have a bigger impact on sensor sensitivity than baseline and lead to change in both slope and correlation of regression models. The decline in performance was significant, reaching 40% decline in range after 2 months and 70% decline after 4 months. Decline in sensor responsiveness over time also led to decline in inter-device variation. Increase

in temperature and humidity led to increase in sensor inter-device variation in baseline and sensitivity. NO_x and O₃ pollutant models were most impacted by drift and interferences. The performance of calibration models declined when the devices were moved to a new location with the magnitude of decline impacted by the relative range of pollutant concentrations at calibration site compared to the new site.

Three categories of approaches for improving sensor accuracy and reproducibility were tested: nonlinear calibration models, variable transformations and calibration data selection. Nonlinear models, particularly random forest ensemble model and support vector machines, were found to improve the accuracy of regression models for NO_x and PM_{2.5}. A new approach was developed, combining variable transformation with training two separate models for background and local pollution components. The approach was found to have higher accuracy, particularly at new deployment sites. Model reproducibility, ability to adjust for multiple combinations of interferences and ability to resolve sites was improved when devices were calibrated at multiple sites, particularly sites with higher pollutant levels. However, incorporation of long-term historical data didn't significantly improve model performance.

Analysis showed that both short-term and long-term temporal patterns, including the variation in peak magnitude, timing and frequency could be resolved and compared at different sites. Traffic-related pollution was highest and most divergent across sites during the morning rush hour. While absolute pollutant levels were indicative of average traffic volume at a site, pollutant temporal transients (i.e. variation in peak timing, frequency, amplitude) revealed insights on local pollution sources. The accuracy of temporal pattern analysis was higher at sites with high and medium traffic. Sensors could also detect changes in CO-to-NO_x ratios, which varied spatially and temporally suggesting that changes in vehicle fleet composition can be detected by AirSENCE devices.

The spatial variability in pollutant level magnitude could be detected by sensor array devices, consistent with those previously identified through Land Use Regression models. The spatial variation in traffic-related pollutants was most significant in the morning and on weekdays, suggesting those are the best time to detect differences between sites. Wind from west was associated with largest difference between sites for NO_x and CO. The sensor device CO model showed the best performance in resolving spatial variation, while NO_x and CO₂ models showed best performance at sites with low and medium traffic levels. This was caused by decline in accuracy at high concentration of pollutant due to sensor

response nonlinearity. The importance of nonlinearity was confirmed by the positive impact of variable transformation on model performance.

The analysis of multi-pollutant patterns supports the analysis of pollution sources. The comparison of the degree of correlation between pollutants at different sites helped reveal source composition patterns. Background subtraction helped further emphasize the differences and rank sites in terms of traffic-related pollution, as indicated by the comparison of spatial variation in local NO_x contribution compared to that of background contribution. NO_x was found to have the higher local contribution component on average across sites, followed by CO and CO₂. Wind direction was also found to be helpful for understanding the effect of local pollution sources. Weekly temporal variation pattern was more variable across sites compared to daily pattern, suggesting that it might be a better metric for site comparison. Raw sensor data was found to be sufficient for detecting differences in source composition variation. This is particularly promising, since significant amount of time and costs can be saved if sensors can be directly useful without calibration. Comparative analysis of temporal patterns of different sensors revealed changes in fleet composition over time.

In summary, the results described above are promising with respect to performance and applications of sensor array devices for monitoring air pollution levels, composition and sources. The data collected by AirSENCE devices enabled monitoring of more air contaminants at higher temporal and spatial resolution compared to traditional approaches. It also provided unique insights into the pollution sources and variation of air pollution in space and time. However, the analysis of the impact of environmental factors and inter-sensor variation has revealed key performance limitations that are important for developing calibration protocols for low-cost sensor devices. For instance, the analysis of the impact of temperature and humidity as well as interferences on sensor performance have suggested novel methods that would help flag faulty measurements and correct sensor response. Key sensor response interferences were identified, and this result will be useful for accessing the level of measurement uncertainty in future work. Another key conclusion of this work is the importance of validation of sensor devices at a range of locations with different pollutant levels and sources. The impact of calibration dataset was explored, and the insights derived from analysis of sensor response to different combinations of interferences will be key for designing sensor maintenance protocols in future studies. A new protocol was proposed for calibration of low-cost sensor devices to ensure generalizable calibration models and better performance across sites. A new method was also developed for evaluating sensor drift, revealing new insights into ideal re-calibration frequency and drift correction

methods. The comprehensive comparison of variable transforms and nonlinear approaches has suggested that some approaches (Support Vector Machine Regression as well as square and exponential variable transforms) clearly outperform others and should be the focus of future studies. Furthermore, the comparison of methods for deriving Air Quality Health Index indicated higher accuracy of direct regression calibration approach – an important insight for design of future studies of air quality indices derived using low-cost sensor measurements. The temporal sensor response patterns revealed unique insights into the differences in level and characteristics of sources such as traffic across sites. The results suggested that background subtraction and analysis of the peak magnitude, timing and frequency would be useful for spatial pollution pattern analysis in future sensor network studies. While some challenges still remain, inroads in tackling them have been made and the results have indicated new opportunities for using the pollution sensor data to improve the design our cities as well as reduce the impact of air pollution on our health and wellbeing (Figure 72).

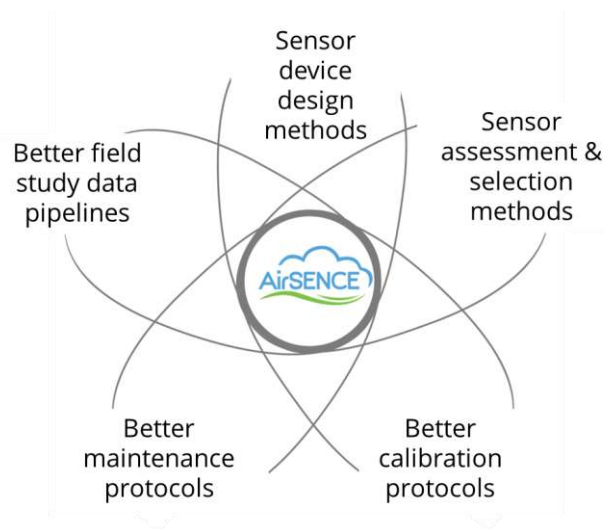


Figure 72. Summary of the key contributions of the thesis to the field of research.

6 Future Work

While many major challenges and opportunities associated with low-cost sensor devices have been addressed in this work, several key opportunities remain. The field of air quality monitoring has matured significantly over the past 10 years. Air pollution sensor technology is evolving quickly. New sensors and air quality monitoring devices have been proposed and adapted into sensor networks. New sensors have been introduced, while others became obsolete. It will be key for researchers and industry alike to continue testing new technologies and develop better calibration protocols that would allow for the calibration models to be transferred across devices (Fang and Bate 2017). Better methods for control of inter-device variation will be important for development of generalizable calibration models. Better calibration protocols will also need to be developed to generate calibration models that perform well across sites with different types and levels of air contaminants as well as across a range of environmental conditions. As this work indicated, calibration of AirSENCE devices using data from a combination of different locations leads to improved performance. However, as the size of sensor network grows, the costs and time required to conduct such calibrations becomes unsustainable. Oversampling is one approach for overcoming this challenge (Balzano and Nowak 2007). Alternative approaches that make use of sensor nodes mounted on public transit or other moving platforms is another solution for acquiring large and diverse calibration datasets quickly (Hasenfratz, Saukh, and Thiele 2012). Another key opportunity for further research is development of correction factors for sensor interferences and drift. A more robust system is needed for flagging interferences in sensor data, categorizing data according to the level of uncertainty as well as automatically adjusting it in real time. Such metadata would also be essential for prioritizing device maintenance and quality control. Finally, improved algorithms for identifying differences between sites and tracking the variation in pollution sources across space and time will be important as low-cost air quality sensor technologies and networks continue to evolve. Fusion of diverse datasets (i.e. satellite data, traffic data and weather data) will be key for the development of these algorithms.

There are 3 major areas of research will continue to be important as the sensors and the air quality monitoring techniques continue to evolve: 1) Improving the hardware, software and data infrastructure required to enable high density measurements, 2) Optimizing frameworks for tracking and optimizing context-specific performance of AirSENCE devices, and 3) Developing new protocols for enabling key emerging applications for tracking pollution sources and exposure.

First, higher density sensor networks will lead to new opportunities. The existing static sensor networks will be supplemented with mobile peer-to-peer networks, where sensors are mounted on vehicles, bicycles and eventually worn by people (Figure 73). The work for supplementing static sensor networks with mobile networks is already in progress at SOCAAR lab, where AirSENCE devices have been installed on two TTC streetcars (Mitchell 2018). AirSENCE technology has been licensed to AUG Signals and is in preparation for commercial launch.

Portable air quality sensor measures pollution levels

AllerGen INVESTIGATORS-INVENTORS:

Drs Greg Evans and Jeff Brook, University of Toronto



Drs Evans and Brook are developing AirSENCE (Air SENSOR for Chemicals in the Environment), a novel technology for monitoring air quality. The device uses a panel of sensors to gauge the levels of five common air pollutants: nitrogen oxides, ozone, carbon monoxide, carbon dioxide, and particulate matter. The portable sensor can be mounted outdoors or indoors, and produces air quality data that can be instantly accessed through a smart phone.

Following a successful pilot test of the device during the 2015 Pan Am Games, the researchers are now conducting a larger-scale evaluation of AirSENCE in Beijing, China and expanding its use in the Greater Toronto Area. By Fall 2017, AirSENCE devices will be installed along the King Street and Yonge Street corridors in Toronto, and in the neighbouring cities of Guelph and Mississauga on lamp posts and utility poles.

In addition to research support from AllerGen since 2007 and \$221,000 in partner contributions secured in 2016-17, the team has financial backing from the Ontario-China Research and Innovation Fund, and is collaborating with AUG Signals, a Canadian company.

Photo: SOCAAR/AirSENCE; technicians mount an AirSENCE device on a King Street (Toronto) streetcar.

Figure 73. AirSENCE devices installed on TTC streetcars for monitoring ozone, carbon monoxide, carbon dioxide, and PM_{2.5} along King Street routes¹¹.

However, for the sensor networks to become truly dense and ubiquitous, sensors need to become more power-efficient and smaller. Some progress in this area has already been made both by academic (Deng et al. 2016) and industrial teams (Figaro 2017). Smartphones containing air pollution sensors have been announced (Jones 2018). Innovative battery-free systems based on radio frequency energy harvesting have been proposed (Tran, Dang, and Chung 2017).

Second, as the sensors become more ubiquitous, a system would need to be developed to automatically track the quality of data and model output. As shown earlier in this document, sensor performance can vary significantly depending on how the sensor device was calibrated as well as the types of nearby

¹¹ <http://annualreports.allergen-nce.ca/ar2016/#c8>

sources and environmental factors. The development of a distributed ledger system for tracking the device, calibration, model and deployment parameters will become key to ensuring the quality of sensor device output moving forward. One method to implement such a system is an open database of sensor response under a range of environmental conditions and source profiles. Such database would serve as a platform for researchers and citizen scientists to upload their own documented datasets. A system for documentation and automatic curation and quality control as well as open hardware and software infrastructure would be essential for helping us establish such a database and steps in the right direction have already been taken (Kotsev et al. 2016). Establishing a system for data governance will be a key need moving forward. The open database would be very useful for developing classification model libraries (i.e. fully reproducible model specifications) that would identify pollution sources and predict pollution variation in the future. These model libraries would become more accurate and generalizable as more sample pollution data is uploaded in the database and would enable people across the world to develop tools and APIs using these data sources. Having reliable open datasets and models would be invaluable to help develop better policies on urban development, pollution source control and climate change mediation. We would also be able to use them to test the outcomes of small-scale interventions (i.e. no-idling laws, electric and biofuel vehicles, new city designs).

Finally, new sensor-specific data analysis techniques should be developed that go beyond improving the sensor performance and support new applications. For example, the analysis of emission factors, source apportionment, exposure assessment and tracking of air pollution interventions are some of the emerging areas of interest. Some recent studies have been published exploring the use of low-cost sensors for evaluating emission factors (Johnson et al. 2016) and deriving exposure estimates (Castell et al. 2017). An ongoing challenge in exposure assessment is that the level of the concentration recorded by a sensor at a specific point in time may not reflect the duration of exposure related to that standard or benchmark (e.g. daily or annual average benchmarks). A new framework for benchmarking data collected by low cost sensors will need to be established.

7 References

- Aberer, Karl, Saket Sathe, Dipanjan Chakraborty, Alcherio Martinoli, Guillermo Barrenetxea, Boi Faltings, and Lothar Thiele. 2010. "OpenSense: Open Community Driven Sensing of Environment Karl." *Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming - GIS '10*, 39–42. doi:10.1145/1878500.1878509.
- Abidin, M. Z., Arnis Asmat, and M. N. Hamidon. 2013. "Identification of Initial Drift in Semiconductor Gas Sensors Caused by Temperature Variation." *Proceedings - 2013 IEEE 9th International Colloquium on Signal Processing and Its Applications, CSPA 2013*, 285–88. doi:10.1109/CSPA.2013.6530057.
- Aguiar, Erick Frederico Kill, Henrique Llacer Roig, Lúis Henrique Mancini, and Eduardo Neiva Caetano Botelho de Carvalho. 2015. "Low-Cost Sensors Calibration for Monitoring Air Quality in the Federal District—Brazil." *Journal of Environmental Protection* 6 (2): 173–89. doi:10.4236/jep.2015.62019.
- Alwadi, Mohammad, and Girija Chetty. 2015. "Energy Efficient Data Mining Scheme for High Dimensional Data." *Procedia Computer Science* 46 (Icict 2014): 483–90. doi:10.1016/j.procs.2015.02.047.
- Araki, Atsuko, Toshio Kawai, Yoko Eitaki, Ayako Kanazawa, Kanehisa Morimoto, Kunio Nakayama, Eiji Shibata, et al. 2010. "Relationship between Selected Indoor Volatile Organic Compounds, so-Called Microbial VOC, and the Prevalence of Mucous Membrane Symptoms in Single Family Homes." *Science of The Total Environment* 408 (10): 2208–15. doi:10.1016/j.scitotenv.2010.02.012.
- Arfire, Adrian, Ali Marjovi, and Alcherio Martinoli. 2016. "Enhancing Measurement Quality through Active Sampling in Mobile Air Quality Monitoring Sensor Networks." In *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, 1022–27. IEEE. doi:10.1109/AIM.2016.7576904.
- ASTM International. 2016. "ASTM International - Standards Worldwide." http://www.astm.org/SNEWS/MJ_2009/nelson_mj09.html.
- Balzano, Laura, and Robert Nowak. 2007. "Blind Calibration of Sensor Networks." *Proceedings of the 6th International Conference on Information Processing in Sensor Networks - IPSN '07*, 79–88. doi:10.1109/IPSN.2007.4379667.
- Barakeh, Zaher Al, Philippe Breuil, Nathalie Redon, Christophe Pijolat, Nadine Locoge, and Jean Paul Viricelle. 2017. "Development of a Normalized Multi-Sensors System for Low Cost on-Line Atmospheric Pollution Detection." *Sensors and Actuators, B: Chemical* 241. Elsevier B.V.: 1235–43. doi:10.1016/j.snb.2016.10.006.
- Barck, C., J. Lundahl, G. Halldén, and G. Bylin. 2005. "Brief Exposures to NO₂ Augment the Allergic

- Inflammation in Asthmatics." *Environmental Research* 97 (1): 58–66.
doi:10.1016/j.envres.2004.02.009.
- Barsan, N., D. Koziej, and U. Weimar. 2007. "Metal Oxide-Based Gas Sensor Research: How To?" *Sensors and Actuators, B: Chemical* 121 (1): 18–35. doi:10.1016/j.snb.2006.09.047.
- Becker, Muhlberger, Braunmuhl, Muller, Ziemann, and Hechtenberg. 2000. "Air Pollution Monitoring Using Tin-Oxide-Based Microreactor Systems." *Sensors and Actuators, B: Chemical* 69 (1): 108–19. doi:10.1016/S0925-4005(00)00516-5.
- Borge, Rafael, Adolfo Narros, Begoña Artíñano, Carlos Yagüe, Francisco Javier Gómez-Moreno, David de la Paz, Carlos Román-Cascón, et al. 2016. "Assessment of Microscale Spatio-Temporal Variation of Air Pollution at an Urban Hotspot in Madrid (Spain) through an Extensive Field Campaign." *Atmospheric Environment* 140: 432–45. doi:10.1016/j.atmosenv.2016.06.020.
- Boubrima, Ahmed, Walid Bechkit, and Herve Rivano. 2016. "Error-Bounded Air Quality Mapping Using Wireless Sensor Networks." *2016 IEEE 41st Conference on Local Computer Networks (LCN)*, 380–88. doi:10.1109/LCN.2016.66.
- Budde, Matthias, Mathias Busse, and Michael Beigl. 2012. "Investigating the Use of Commodity Dust Sensors for the Embedded Measurement of Particulate Matter." *2012 Ninth International Conference on Networked Sensing Systems (INSS)*, 1–4. doi:10.1109/INSS.2012.6240545.
- Budde, Matthias, Marcel Köpke, and Michael Beigl. 2015. "Robust in-Situ Data Reconstruction from Poisson Noise for Low-Cost, Mobile, Non-Expert Environmental Sensing." *Proceedings of the 2015 ACM International Symposium on Wearable Computers - ISWC '15*, 179–82. doi:10.1145/2802083.2808406.
- Budde, Matthias, Rayan El Masri, Till Riedel, and Michael Beigl. 2013. "Enabling Low-Cost Particulate Matter Measurement for Participatory Sensing Scenarios." *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia - MUM '13*, no. Mum: 1–10. doi:10.1145/2541831.2541859.
- Buka, Irena, Samuel Koranteng, Mb Chb, Alvaro R Osornio-Vargas, and Md Phd. 2006. "The Effects of Air Pollution on the Health of Children." *Paediatr Child Health* 1111 (8): 513–16.
- Caron, Redon, Thevenet, Hanoune, and Coddeville. 2016. "Performances and Limitations of Electronic Gas Sensors to Investigate an Indoor Air Quality Event." *Building and Environment* 107. Elsevier Ltd: 19–28. doi:10.1016/j.buildenv.2016.07.006.
- Carton, L. 2015. "FILLING THE FEEDBACK GAP OF PLACE-RELATED 'EXTERNALITIES' IN SMART CITIES: Empowering Citizen-Sensor-Networks for Participatory Monitoring and Planning for a Responsible Distribution of Urban Air Quality Linda," 1–22.

- Castell, Nuria, Franck R. Dauge, Philipp Schneider, Matthias Vogt, Uri Lerner, Barak Fishbain, David Broday, and Alena Bartonova. 2017. "Can Commercial Low-Cost Sensor Platforms Contribute to Air Quality Monitoring and Exposure Estimates?" *Environment International* 99. The Authors: 293–302. doi:10.1016/j.envint.2016.12.007.
- Chaiyboun, A., R. Traute, T. Haas, O. Kieseewetter, and T. Doll. 2007. "A Logarithmic Multi-Parameter Model Using Gas Sensor Main and Cross Sensitivities to Estimate Gas Concentrations in a Gas Mixture for SnO₂ Gas Sensors." *Sensors and Actuators, B: Chemical* 123 (2): 1064–70. doi:10.1016/j.snb.2006.11.012.
- Chang, Ta-Yuan, Shen-Ju Lin, Ruei-Hao Shie, Shih-Wei Tsai, Hui-Tsung Hsu, Ching-Tsan Tsai, Hsien-Wen Kuo, Chow-Feng Chiang, and Jim-Shoung Lai. 2010. "Characterization of Volatile Organic Compounds in the Vicinity of an Optoelectronics Industrial Park in Taiwan." *Journal of the Air & Waste Management Association (1995)* 60 (1): 55–62. <http://www.ncbi.nlm.nih.gov/pubmed/20102035>.
- Choi, Kwang-Yong, Joon-Shik Park, Kwang-Bum Park, Hyun Jae Kim, Hyo-Derk Park, and Seong-Dong Kim. 2010. "Low Power Micro-Gas Sensors Using Mixed SnO₂ Nanoparticles and MWCNTs to Detect NO₂, NH₃, and Xylene Gases for Ubiquitous Sensor Network Applications." *Sensors and Actuators B: Chemical* 150 (1): 65–72. doi:10.1016/j.snb.2010.07.041.
- Cindoruk, S. Siddik, and Yücel Tasdemir. 2010. "Dynamics of Atmospheric Polychlorinated Biphenyls (PCBs): Concentrations, Patterns, Partitioning, and Dry Deposition Level Estimations in a Residential Site of Turkey." *Environmental Monitoring and Assessment* 162 (1–4). Springer Netherlands: 67–80. doi:10.1007/s10661-009-0776-1.
- Corrêa, Sergio Machado, Graciela Arbilla, Eduardo Monteiro Martins, Simone Lorena Quitério, Claudinei de Souza Guimarães, and Luciana Vanni Gatti. 2010. "Five Years of Formaldehyde and Acetaldehyde Monitoring in the Rio de Janeiro Downtown Area – Brazil." *Atmospheric Environment* 44 (19): 2302–8. doi:10.1016/j.atmosenv.2010.03.043.
- Delpha, Claude, Maryam Siadat, and Martine Lumbreras. 1999. "Humidity Dependence of a TGS Gas Sensor Array in an Air-Conditioned Atmosphere." *Sensors and Actuators, B: Chemical* 59 (2): 255–59. doi:10.1016/S0925-4005(99)00230-0.
- Deng, Yue, Cheng Chen, Xiaojun Xian, Francis Tsow, Gaurav Verma, Rob McConnell, Scott Fruin, Nongjian Tao, and Erica S. Forzani. 2016. "A Novel Wireless Wearable Volatile Organic Compound (VOC) Monitoring Device with Disposable Sensors." *Sensors (Switzerland)* 16 (12). doi:10.3390/s16122060.
- Devarakonda, Srinivas, Parveen Sevusu, Hongzhang Liu, Ruilin Liu, Liviu Iftode, and Badri Nath. 2013.

- "Real-Time Air Quality Monitoring through Mobile Sensing in Metropolitan Areas." *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing - UrbComp '13*, 1. doi:10.1145/2505821.2505834.
- DeVito, Saverio, Anna Castaldo, Fausta Loffredo, Ettore Massera, Tiziana Polichetti, Ivana Nasti, Paolo Vacca, Luigi Quercia, and Girolamo Di Francia. 2007. "Gas Concentration Estimation in Ternary Mixtures with Room Temperature Operating Sensor Array Using Tapped Delay Architectures." *Sensors and Actuators, B: Chemical* 124 (2): 309–16. doi:10.1016/j.snb.2006.12.039.
- DeVito, Massera, Piga, Martinotto, and Di Francia. 2008. "On Field Calibration of an Electronic Nose for Benzene Estimation in an Urban Pollution Monitoring Scenario." *Sensors and Actuators, B: Chemical* 129 (2): 750–57. doi:10.1016/j.snb.2007.09.060.
- DeVito, Piga, Martinotto, and Di Francia. 2009. "CO, NO₂ and NO_x Urban Pollution Monitoring with on-Field Calibrated Electronic Nose by Automatic Bayesian Regularization." *Sensors and Actuators, B: Chemical* 143 (1): 182–91. doi:10.1016/j.snb.2009.08.041.
- Du, Yi, Cuixia Ma, Chao Wu, Xiaowei Xu, Yike Guo, Yuanchun Zhou, and Jianhui Li. 2017. "A Visual Analytics Approach for Station-Based Air Quality Data." *Sensors (Switzerland)* 17 (1): 1–17. doi:10.3390/s17010030.
- Esposito, E., S. De Vito, M. Salvato, V. Bright, R. L. Jones, and O. Popoola. 2016. "Dynamic Neural Network Architectures for on Field Stochastic Calibration of Indicative Low Cost Air Quality Sensing Systems." *Sensors and Actuators, B: Chemical* 231. Elsevier B.V.: 701–13. doi:10.1016/j.snb.2016.03.038.
- Fang, Xinwei, and Iain John Bate. 2017. "Issues of Using Wireless Sensor Network to Monitor Urban Air Quality." ACM. <http://eprints.whiterose.ac.uk/125046/>.
- Fernandez, Luis, S. Guney, A. Gutierrez-Galvez, and S. Marco. 2016. "Calibration Transfer in Temperature Modulated Gas Sensor Arrays." *Sensors and Actuators, B: Chemical* 231. Elsevier B.V.: 276–84. doi:10.1016/j.snb.2016.02.131.
- Fernandez, Marco, and Gutierrez-Galvez. 2015. "Robustness to Sensor Damage of a Highly Redundant Gas Sensor Array." *Sensors and Actuators, B: Chemical* 218. Elsevier B.V.: 296–302. doi:10.1016/j.snb.2015.04.096.
- Figaro. 2017. "TGS8100 - Featured Products - FIGARO Engineering Inc." <http://www.figaro.co.jp/en/product/feature/tgs8100.html>.
- Fine, George F., Leon M. Cavanagh, Ayo Afonja, and Russell Binions. 2010. "Metal Oxide Semi-Conductor Gas Sensors in Environmental Monitoring." *Sensors* 10 (6): 5469–5502. doi:10.3390/s100605469.
- Fischer, Joseph C. von, Daniel Cooley, Sam Chamberlain, Adam Gaylord, Claire J. Griebenow, Steven P.

- Hamburg, Jessica Salo, Russ Schumacher, David Theobald, and Jay Ham. 2017. "Rapid, Vehicle-Based Identification of Location and Magnitude of Urban Natural Gas Pipeline Leaks." *Environmental Science & Technology*, acs.est.6b06095. doi:10.1021/acs.est.6b06095.
- Fonollosa, Jordi, Luis Fernández, Ramón Huerta, Agustín Gutiérrez-Gálvez, and Santiago Marco. 2013. "Temperature Optimization of Metal Oxide Sensor Arrays Using Mutual Information." *Sensors and Actuators, B: Chemical* 187. Elsevier B.V.: 331–39. doi:10.1016/j.snb.2012.12.026.
- Fonollosa, Jordi, Alexander Vergara, Ramón Huerta, and Santiago Marco. 2014. "Estimation of the Limit of Detection Using Information Theory Measures." *Analytica Chimica Acta* 810. Elsevier B.V.: 1–9. doi:10.1016/j.aca.2013.10.030.
- Fonollosa, L. Fernández, A. Gutiérrez-Gálvez, R. Huerta, and S. Marco. 2016. "Calibration Transfer and Drift Counteraction in Chemical Sensor Arrays Using Direct Standardization." *Sensors and Actuators, B: Chemical* 236. Elsevier B.V.: 1044–53. doi:10.1016/j.snb.2016.05.089.
- Fonollosa, Vergara, and Huerta. 2013. "Algorithmic Mitigation of Sensor Failure: Is Sensor Replacement Really Necessary?" *Sensors and Actuators, B: Chemical* 183. Elsevier B.V.: 211–21. doi:10.1016/j.snb.2013.03.034.
- Fuertes, Walter, Diego Carrera, Cesar Villacis, Theofilos Toulkeridis, Fernando Galarraga, Edgar Torres, and Hernan Aules. 2016. "Distributed System as Internet of Things for a New Low-Cost, Air Pollution Wireless Monitoring on Real Time." *Proceedings - 2015 IEEE/ACM 19th International Symposium on Distributed Simulation and Real Time Applications, DS-RT 2015*, 58–67. doi:10.1109/DS-RT.2015.28.
- Fuzzi, S., U. Baltensperger, K. Carslaw, S. Decesari, H. Denier Van Der Gon, M. C. Facchini, D. Fowler, et al. 2015. "Particulate Matter, Air Quality and Climate: Lessons Learned and Future Needs." *Atmospheric Chemistry and Physics* 15 (14): 8217–99. doi:10.5194/acp-15-8217-2015.
- Gao, Meiling, Junji Cao, and Edmund Seto. 2015. "A Distributed Network of Low-Cost Continuous Reading Sensors to Measure Spatiotemporal Variations of PM_{2.5} in Xi'an, China." *Environmental Pollution* 199. Elsevier Ltd: 56–65. doi:10.1016/j.envpol.2015.01.013.
- Gardner, J. W., P. Boilot, and E. L. Hines. 2005. "Enhancing Electronic Nose Performance by Sensor Selection Using a New Integer-Based Genetic Algorithm Approach." *Sensors and Actuators, B: Chemical* 106 (1 SPEC. ISS.): 114–21. doi:10.1016/j.snb.2004.05.043.
- Gerboles, Michel, and Daniela Buzica. 2009. *EVALUATION OF MICRO-SENSORS TO MONITOR OZONE IN AMBIENT AIR*. doi:10.2788/5978.
- Gordon, David, Mary Shaw, Daniel Dorling, and George Davey Smith. 1999. *Inequalities in Health : The Evidence ; the Evidence Presented to the Independent Inquiry into Inequalities in Health, Chaired by*

- Sir Donald Acheson*. Policy Press. https://books.google.it/books?id=YuRdGdP5MXEC&pg=RA1-PT87&lpg=RA1-PT87&dq=air+pollution+children,+the+elderly,+and+people+with+disabilities&source=bl&ots=sNltK5Uo5f&sig=kebZ3fiuyp_ZFrGk-1gVkv46qq&hl=it&sa=X&ved=0ahUKEwjpsa2bwoPWAhUERhQKH85BiAQ6AEIdzAJ#v.1gVkv46qq&hl=it&sa=X&ved=0ahUKEwjpsa2bwoPWAhUERhQKH85BiAQ6AEIdzAJ#v.
- Gualdrón, O., J. Brezmes, E. Llobet, A. Amari, X. Vilanova, B. Bouchikhi, and X. Correig. 2007. "Variable Selection for Support Vector Machine Based Multisensor Systems." *Sensors and Actuators, B: Chemical* 122 (1): 259–68. doi:10.1016/j.snb.2006.05.029.
- Gualdrón, O., E. Llobet, J. Brezmes, X. Vilanova, and X. Correig. 2006. "Coupling Fast Variable Selection Methods to Neural Network-Based Classifiers: Application to Multisensor Systems." *Sensors and Actuators, B: Chemical* 114 (1): 522–29. doi:10.1016/j.snb.2005.04.046.
- Hamada, Kaoru, Carroll-Aan Goldsmith, Alejandra Goldman, and Lester Kobzik. 2000. "Resistance of Very Young Mice to Inhaled Allergen Sensitization Is Overcome by Coexposure to an Air-Pollutant Aerosol." *American Journal of Respiratory and Critical Care Medicine* 161 (4): 1285–93. doi:10.1164/ajrccm.161.4.9906137.
- Hasenfratz, David, Olga Saukh, and Lothar Thiele. 2012. "On-the-Fly Calibration of Low-Cost Gas Sensors." *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7158 LNCS: 228–44. doi:10.1007/978-3-642-28169-3_15.
- Hasenfratz, David, Olga Saukh, Christoph Walser, Christoph Hueglin, Martin Fierz, Tabita Arn, Jan Beutel, and Lothar Thiele. 2015. "Deriving High-Resolution Urban Air Pollution Maps Using Mobile Sensor Nodes." *Pervasive and Mobile Computing* 16 (PB). Elsevier B.V.: 268–85. doi:10.1016/j.pmcj.2014.11.008.
- Haugen, John Erik, Oliver Tomic, and Knut Kvaal. 2000. "A Calibration Method for Handling the Temporal Drift of Solid State Gas-Sensors." *Analytica Chimica Acta* 407 (1–2): 23–39. doi:10.1016/S0003-2670(99)00784-9.
- He, Jie, Liyuan Xu, Peng Wang, and Qin Wang. 2017. "A High Precise E-Nose for Daily Indoor Air Quality Monitoring in Living Environment." *Integration, the VLSI Journal* 58 (xxxx). Elsevier: 286–94. doi:10.1016/j.vlsi.2016.12.010.
- Helwig, Muller, Sberveglieri, and Eickhoff. 2009. "On the Low-Temperature Response of Semiconductor Gas Sensors." *Journal of Sensors* 2009. doi:10.1155/2009/620720.
- Hirobayashi, Shigeki, Mohammed Afrose Kadir, Toshio Yoshizawa, and Tatsuo Yamabuchi. 2003. "Verification of a Logarithmic Model for Estimation of Gas Concentrations in a Mixture for a Tin

- Oxide Gas Sensor Response." *Sensors and Actuators, B: Chemical* 92 (3): 269–78.
doi:10.1016/S0925-4005(03)00311-3.
- Hirobayashi, Shigeki, Haruhiko Kimura, and Takashi Oyabu. 1999. "Dynamic Model to Estimate the Dependence of Gas Sensor Characteristics on Temperature and Humidity in Environment." *Sensors and Actuators, B: Chemical* 60 (1): 78–82. doi:10.1016/S0925-4005(99)00249-X.
- Holmberg, Martin, Fabrizio a.M. Davide, Corrado Di Natale, Arnaldo D'Amico, Fredrik Winqvist, and Ingemar Lundström. 1997. "Drift Counteraction in Odour Recognition Applications: Lifelong Calibration Method." *Sensors and Actuators B: Chemical* 42 (97): 185–94. doi:10.1016/S0925-4005(97)80335-8.
- Holstius, D. M., A. Pillarisetti, K. R. Smith, and E. Seto. 2014. "Field Calibrations of a Low-Cost Aerosol Sensor at a Regulatory Monitoring Site in California." *Atmospheric Measurement Techniques* 7 (4): 1121–31. doi:10.5194/amt-7-1121-2014.
- Huang, Tailai, Pengfei Jia, Peilin He, Shukai Duan, Jia Yan, and Lidan Wang. 2016. "A Novel Semi-Supervised Method of Electronic Nose for Indoor Pollution Detection Trained by M-S4VMS." *Sensors (Switzerland)* 16 (9). doi:10.3390/s16091462.
- Huerta, Ramon, Thiago Mosqueiro, Jordi Fonollosa, Nikolai F. Rulkov, and Irene Rodriguez-Lujan. 2016. "Online Decorrelation of Humidity and Temperature in Chemical Sensors for Continuous Monitoring." *Chemometrics and Intelligent Laboratory Systems* 157. Elsevier B.V.: 169–76. doi:10.1016/j.chemolab.2016.07.004.
- IEC. 2017. "Welcome to the IEC - International Electrotechnical Commission." <http://www.iec.ch/>.
- ISO. 2005. "ISO/IEC 17025:2005 - General Requirements for the Competence of Testing and Calibration Laboratories." <https://www.iso.org/standard/39883.html>.
- Jerrett, Michael. 2015. "Atmospheric Science: The Death Toll from Air-Pollution Sources." *Nature* 525 (7569): 330–31. doi:10.1038/525330a.
- Jerrett, Michael, David Donaire-Gonzalez, Olalekan Popoola, Roderic Jones, Ronald C. Cohen, Estela Almanza, Audrey de Nazelle, et al. 2017. "Validating Novel Air Pollution Sensors to Improve Exposure Estimates for Epidemiological Analyses and Citizen Science." *Environmental Research* 158 (March). Elsevier Inc.: 286–94. doi:10.1016/j.envres.2017.04.023.
- Jiao, Wan, Gayle Hagler, Ronald Williams, Robert Sharpe, Ryan Brown, Daniel Garver, Robert Judge, et al. 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): 5281–92. doi:10.5194/amt-9-5281-2016.
- Johnson, Karoline K., Michael H. Bergin, Armistead G. Russell, and Gayle S. W. Hagler. 2016. "Using Low

- Cost Sensors to Measure Ambient Particulate Matter Concentrations and On-Road Emissions Factors." *Atmospheric Measurement Techniques Discussions*, no. February: 1–22. doi:10.5194/amt-2015-331.
- Jones. 2018. "Cat S61 Smartphone Will Let You See Through Walls, Test Air Quality." *ValueWalk*.
<http://www.valuewalk.com/2018/02/cat-s61-smartphone-camera/>.
- Jovasevic-Stojanovic, Milena, Alena Bartonova, Dusan Topalovic, Ivan Lazovic, Boris Pokric, and Zoran Ristovski. 2015. "On the Use of Small and Cheaper Sensors and Devices for Indicative Citizen-Based Monitoring of Respirable Particulate Matter." *Environmental Pollution* 206: 696–704.
 doi:10.1016/j.envpol.2015.08.035.
- Kadri, Chaibou, Fengchun Tian, Lei Zhang, Guojun Li, Lijun Dang, and Guorui Li. 2013. "Neural Network Ensembles for Online Gas Concentration Estimation Using an Electronic Nose." *Int. J. Comput. Sci* 10 (2013): 10. <http://www.ijcsi.org/papers/IJCSI-10-2-1-129-135.pdf>.
- Kanan, Sofian M., Oussama M. El-Kadri, Imad A. Abu-Yousef, and Marsha C. Kanan. 2009. "Semiconducting Metal Oxide Based Sensors for Selective Gas Pollutant Detection." *Sensors* 9 (10): 8158–96. doi:10.3390/s91008158.
- Kelly, K. E., J. Whitaker, A. Petty, C. Widmer, A. Dybwad, D. Sleeth, R. Martin, and A. Butterfield. 2017. "Ambient and Laboratory Evaluation of a Low-Cost Particulate Matter Sensor." *Environmental Pollution* 221: 491–500. doi:10.1016/j.envpol.2016.12.039.
- Khatko, V., S. Vallejos, J. Calderer, I. Gracia, C. Cané, E. Llobet, and X. Correig. 2009. "Micro-Machined WO₃-Based Sensors with Improved Characteristics." *Sensors and Actuators B: Chemical* 140 (2): 356–62. doi:10.1016/j.snb.2009.05.020.
- Korotcenkov, G. 2007. "Metal Oxides for Solid-State Gas Sensors: What Determines Our Choice?" *Materials Science and Engineering B: Solid-State Materials for Advanced Technology* 139 (1): 1–23. doi:10.1016/j.mseb.2007.01.044.
- Kortenkamp, Andreas, Thomas Backhaus, and Michael Faust. 2009. "State of the Art Report on Mixture Toxicity." *Report Contract N* (December): 1–391.
 doi:http://ec.europa.eu/environment/chemicals/pdf/report_Mixture%20toxicity.pdf.
- Kotsev, Alexander, Sven Schade, Massimo Craglia, Michel Gerboles, Laurent Spinnelle, and Marco Signorini. 2016. "Next Generation Air Quality Platform: Openness and Interoperability for the Internet of Things." *Sensors (Switzerland)* 16 (3): 1–16. doi:10.3390/s16030403.
- Kumar, Prashant, Lidia Morawska, Claudio Martani, George Biskos, Marina Neophytou, Silvana Di Sabatino, Margaret Bell, Leslie Norford, and Rex Britter. 2015. "The Rise of Low-Cost Sensing for Managing Air Pollution in Cities." *Environment International* 75. Elsevier Ltd: 199–205.

doi:10.1016/j.envint.2014.11.019.

- Kumar, Prashant, Andreas N. Skouloudis, Margaret Bell, Mar Viana, M. Cristina Carotta, George Biskos, and Lidia Morawska. 2016. "Real-Time Sensors for Indoor Air Monitoring and Challenges Ahead in Deploying Them to Urban Buildings." *Science of the Total Environment* 560–561 (April). Elsevier B.V.: 150–59. doi:10.1016/j.scitotenv.2016.04.032.
- Leoni, Cecilia, Petra Pokorná, Jan Hovorka, Mauro Masiol, Jan Topinka, Yongjing Zhao, Kamil Křůmal, Steven Cliff, Pavel Mikuška, and Philip K. Hopke. 2018. "Source Apportionment of Aerosol Particles at a European Air Pollution Hot Spot Using Particle Number Size Distributions and Chemical Composition." *Environmental Pollution* 234 (March). Elsevier: 145–54. doi:10.1016/J.ENVPOL.2017.10.097.
- Lewis, Alastair, and Peter Edwards. 2017. "Validate Personal Air - Pollution Sensors," 1–7.
- Lewis, Lee, Edwards, Shaw, Evans, Moller, Smith, et al. 2016. "Evaluating the Performance of Low Cost Chemical Sensors for Air Pollution Research." *Faraday Discuss.* 189. Royal Society of Chemistry: 85–103. doi:10.1039/C5FD00201J.
- Li, Xiang, Ling Peng, Yuan Hu, Jing Shao, and Tianhe Chi. 2016. "Deep Learning Architecture for Air Quality Predictions." *Environmental Science and Pollution Research* 23 (22). Environmental Science and Pollution Research: 22408–17. doi:10.1007/s11356-016-7812-9.
- Liu, Hang, Renzhi Chu, and Zhenan Tang. 2015. "Metal Oxide Gas Sensor Drift Compensation Using a Two-Dimensional Classifier Ensemble." *Sensors* 15 (5): 10180–93. doi:10.3390/s150510180.
- Liu, Cheng, Liu, Hu, Zhang, and Ning. 2012. "A Survey on Gas Sensing Technology." *Sensors (Switzerland)* 12 (7): 9635–65. doi:10.3390/s120709635.
- Llobet, Eduard, Jesus Brezmes, Oscar Gualdron, Xavier Vilanova, and Xavier Correig. 2004. "Building Parsimonious Fuzzy ARTMAP Models by Variable Selection with a Cascaded Genetic Algorithm: Application to Multisensor Systems for Gas Analysis." *Sensors and Actuators, B: Chemical* 99 (2–3): 267–72. doi:10.1016/j.snb.2003.11.019.
- MacDonnell, M., D. Raymond, M. Wyker, Y. Finster, T. Chang, B. Raymond, M. Temple, et al. 2014. "Mobile Sensors and Applications for Air Pollutants," no. October: 293. http://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=273979&simpleSearch=1&searchAll=mobile+sensors+and+applications+for+air+pollutants.
- Malaver, Alexander, Nunzio Motta, Peter Corke, and Felipe Gonzalez. 2015. "Development and Integration of a Solar Powered Unmanned Aerial Vehicle and a Wireless Sensor Network to Monitor Greenhouse Gases." *Sensors (Basel, Switzerland)* 15 (2): 4072–96. doi:10.3390/s150204072.

- Manikonda, Abhisek, Nadezda Zikova, Philip K. Hopke, and Andrea R. Ferro. 2016. "Laboratory Assessment of Low-Cost PM Monitors." *Journal of Aerosol Science* 102: 29–40. doi:10.1016/j.jaerosci.2016.08.010.
- Masson, N., R. Piedrahita, and M. Hannigan. 2015. "Approach for Quantification of Metal Oxide Type Semiconductor Gas Sensors Used for Ambient Air Quality Monitoring." *Sensors and Actuators, B: Chemical* 208. Elsevier B.V.: 339–45. doi:10.1016/j.snb.2014.11.032.
- McKercher, Grant R., Jennifer A. Salmond, and Jennifer K. Vanos. 2017. "Characteristics and Applications of Small, Portable Gaseous Air Pollution Monitors." *Environmental Pollution* 223. Elsevier Ltd: 102–10. doi:10.1016/j.envpol.2016.12.045.
- Ministry of the Environment and Climate Change. 2017. "Operations Manual for Air Quality Monitoring in Ontario | Ontario.ca." Accessed October 1. <https://www.ontario.ca/document/operations-manual-air-quality-monitoring-ontario-0>.
- Miskell, Georgia, Jennifer Salmond, and David E. Williams. 2017. "Low-Cost Sensors and Crowd-Sourced Data: Observations of Siting Impacts on a Network of Air-Quality Instruments." *Science of the Total Environment* 575. Elsevier B.V.: 1119–29. doi:10.1016/j.scitotenv.2016.09.177.
- Mitchell. 2018. "Crunching the Numbers on Toronto's King Street Transit Pilot - U of T Engineering News." *UofT Engineering News*. <http://news.engineering.utoronto.ca/crunching-numbers-torontos-king-street-transit-pilot/>.
- Moltchanov, Sharon, Ilan Levy, Yael Etzion, Uri Lerner, David M Broday, and Barak Fishbain. 2015. "On the Feasibility of Measuring Urban Air Pollution by Wireless Distributed Sensor Networks." *Science of the Total Environment, The* 502: 537–47. doi:10.1016/j.scitotenv.2014.09.059.
- Monroy, Javier G., Javier Gonzalez-Jirenez, and Jose Luis Blanco. 2012. "Overcoming the Slow Recovery of MOX Gas Sensors through a System Modeling Approach." *Sensors (Switzerland)* 12 (10): 13664–80. doi:10.3390/s121013664.
- Muezzinoglu, Mehmet K., Alexander Vergara, Ramon Huerta, Nikolai Rulkov, Mikhail I. Rabinovich, Al Selverston, and Henry D I Abarbanel. 2009. "Acceleration of Chemo-Sensory Information Processing Using Transient Features." *Sensors and Actuators, B: Chemical* 137 (2): 507–12. doi:10.1016/j.snb.2008.10.065.
- Natale, Corrado Di, Eugenio Martinelli, and Arnaldo D'Amico. 2002. "Counteraction of Environmental Disturbances of Electronic Nose Data by Independent Component Analysis." *Sensors and Actuators, B: Chemical* 82 (2–3): 158–65. doi:10.1016/S0925-4005(01)01001-2.
- Nowotny, Thomas, Amalia Z. Berna, Russell Binions, and Stephen Trowell. 2013. "Optimal Feature Selection for Classifying a Large Set of Chemicals Using Metal Oxide Sensors." *Sensors and*

- Actuators, B: Chemical* 187. Elsevier B.V.: 471–80. doi:10.1016/j.snb.2013.01.088.
- Olivares, G., and S. Edwards. 2015. “The Outdoor Dust Information Node (ODIN) – Development and Performance Assessment of a Low Cost Ambient Dust Sensor.” *Atmospheric Measurement Techniques Discussions* 8 (7): 7511–33. doi:10.5194/amtd-8-7511-2015.
- Ou, Jiamin, Junyu Zheng, Zibing Yuan, Dabo Guan, Zhijiong Huang, Fei Yu, Min Shao, and Peter K.K. Louie. 2018. “Reconciling Discrepancies in the Source Characterization of VOCs between Emission Inventories and Receptor Modeling.” *Science of The Total Environment* 628–629 (July). Elsevier: 697–706. doi:10.1016/J.SCITOTENV.2018.02.102.
- Pardo, M., and G. Sberveglieri. 2007. “Comparing the Performance of Different Features in Sensor Arrays.” *Sensors and Actuators, B: Chemical* 123 (1): 437–43. doi:10.1016/j.snb.2006.09.041.
- Penza, M., R. Rossi, M. Alvisi, G. Cassano, M.A. Signore, E. Serra, and R. Giorgi. 2008. “Pt- and Pd-Nanoclusters Functionalized Carbon Nanotubes Networked Films for Sub-Ppm Gas Sensors.” *Sensors and Actuators B: Chemical* 135 (1): 289–97. doi:10.1016/j.snb.2008.08.024.
- Penza, Suriano, Villani, Spinelle, and Gerboles. 2014. “Towards Air Quality Indices in Smart Cities by Calibrated Low-Cost Sensors Applied to Networks.” *Proceedings of IEEE Sensors 2014–Decem* (December): 2012–17. doi:10.1109/ICSENS.2014.6985429.
- Peterson, Philip J D, Amrita Aujla, Kirsty H Grant, Alex G Brundle, Martin R Thompson, Josh Vande Hey, and Roland J Leigh. 2017. “Practical Use of Metal Oxide Semiconductor Gas Sensors for Measuring Nitrogen Dioxide and Ozone in Urban Environments.” doi:10.3390/s17071653.
- Piedrahita, R., Y. Xiang, N. Masson, J. Ortega, A. Collier, Y. Jiang, K. Li, et al. 2014. “The next Generation of Low-Cost Personal Air Quality Sensors for Quantitative Exposure Monitoring.” *Atmospheric Measurement Techniques* 7 (10): 3325–36. doi:10.5194/amt-7-3325-2014.
- Plaia, A., F. Di Salvo, M. Ruggieri, and G. Agró. 2013. “A Multisite-Multipollutant Air Quality Index.” *Atmospheric Environment* 70. Elsevier Ltd: 387–91. doi:10.1016/j.atmosenv.2013.01.028.
- Popoola, Olalekan A M, Gregor B. Stewart, Mohammed I. Mead, and Roderic L. Jones. 2016. “Development of a Baseline-Temperature Correction Methodology for Electrochemical Sensors and Its Implications for Long-Term Stability.” *Atmospheric Environment* 147. Elsevier Ltd: 330–43. doi:10.1016/j.atmosenv.2016.10.024.
- Ramadan, A. A. 2010. “Air Quality Assessment in Southern Kuwait Using Diffusive Passive Samplers.” *Environmental Monitoring and Assessment* 160 (1–4): 413–23. doi:10.1007/s10661-008-0705-8.
- Reis, Stefan, Edmund Seto, Amanda Northcross, Nigel W.T. Quinn, Matteo Convertino, Rod L. Jones, Holger R. Maier, et al. 2015. “Integrating Modelling and Smart Sensors for Environmental and Human Health.” *Environmental Modelling and Software* 74: 238–46.

doi:10.1016/j.envsoft.2015.06.003.

Robinson, Johanna A., David Kocman, Milena Horvat, Yael Etzion, Fadi Kizel, Yaela N. Golumbic, Tamar Yacobi, and Dana Draher. 2017. "An Evaluation Tool Kit of Air Quality Micro-Sensing Units." *Science of the Total Environment* 575 (September 2016). Elsevier B.V.: 639–48.

doi:10.1016/j.scitotenv.2016.09.061.

Rodriguez-Lujan, Irene, Jordi Fonollosa, Alexander Vergara, Margie Homer, and Ramon Huerta. 2014. "On the Calibration of Sensor Arrays for Pattern Recognition Using the Minimal Number of Experiments." *Chemometrics and Intelligent Laboratory Systems* 130. Elsevier B.V.: 123–34.

doi:10.1016/j.chemolab.2013.10.012.

Romain, A.C., and J. Nicolas. 2007. "Monitoring Malodours in the Environment with an Electronic Nose : Requirements for the Signal Processing ." In *Biologically Inspired Signal Processing for Chemical Sensing*, SCI 188:121–35.

Salamone, Francesco, Ludovico Danza, Italo Meroni, and Mariacristina Pollastro. 2017. "A Low-Cost Environmental Monitoring System: How to Prevent Systematic Errors in the Design Phase through the Combined Use of Additive Manufacturing and Thermographic Techniques." *Sensors* 17 (4).

doi:10.3390/s17040828.

Sorry, Frédéric, and Martine Lumbreras. 1999. "Gas Composition Determination in an Air Conditioned System Using a Sensor Array: Characterization of Three Different TGS Sensors." *Sensors and Actuators, B: Chemical* 59 (2): 94–99. doi:10.1016/S0925-4005(99)00202-6.

Schutze, Peter Reimann and Andreas. 2008. "Sensor Arrays, Virtual Multisensors, Data Fusion, and Gas Sensor Data Evaluation." *October* 48: 1–32. doi:10.1007/5346.

Seto, E, E Austin, I Novosselov, and M Yost. 2014. "Use of Low-Cost Particle Monitors to Calibrate Traffic-Related Air Pollutant Models in Urban Areas." *7th International Congress on Environmental Modelling and Software, iEMSs 2014* 2: 1009–16.

<http://www.scopus.com/inward/record.url?eid=2-s2.0->

84911877702&partnerID=40&md5=8192eb18e719826d53b3fd67ad244e31.

Shaban, K Bashir, A Kadri, and E Rezk. 2016. "Urban Air Pollution Monitoring System With Forecasting Models." *IEEE Sensors Journal* 16 (8): 2598–2606. doi:10.1109/JSEN.2016.2514378.

Smith, Katie, Peter M Edwards, Mathew John J Evans, James D Lee, Marvin David Shaw, Freya Squires, Shona Wilde, and Alastair C. Lewis. 2017. "Clustering Approaches That Improve the Reproducibility of Low-Cost Air Pollution Sensors." *Faraday Discuss.* doi:10.1039/C7FD00020K.

Snyder, Emily G, Timothy H Watkins, Paul A Solomon, Eben D Thoma, Ronald W Williams, Gayle S W Hagler, David Shelow, David A Hindin, Vasu J Kilaru, and Peter W Preuss. 2013. "The Changing

- Paradigm of Air Pollution Monitoring." *Environmental Science & Technology* 47 (20): 11369–77. doi:10.1021/es4022602.
- Sofowote, Healy, Su, Debosz, Noble, Munoz, Jeong, et al. 2017. "Understanding the PM_{2.5} Imbalance between a Far and near-Road Location: Results of High Temporal Frequency Source Apportionment and Parameterization of Black Carbon." *Atmospheric Environment* Submitted (November). Pergamon. doi:10.1016/j.atmosenv.2017.10.063.
- Sowlat, Mohammad Hossein, Hamed Gharibi, Masud Yunesian, Maryam Tayefeh Mahmoudi, and Saeedeh Lotfi. 2011. "A Novel, Fuzzy-Based Air Quality Index (FAQI) for Air Quality Assessment." *Atmospheric Environment* 45 (12): 2050–59. doi:10.1016/j.atmosenv.2011.01.060.
- Spinelle, Laurent, Michel Gerboles, and Manuel Aleixandre. 2015. "Performance Evaluation of Amperometric Sensors for the Monitoring of O₃ and NO₂ in Ambient Air at Ppb Level." *Procedia Engineering* 120: 480–83. doi:10.1016/j.proeng.2015.08.676.
- Spinelle, Gerboles, and Aleixandre. 2013. *Report of Laboratory and in-Situ Validation of Micro-Sensor Sensor for Monitoring Ambient Air Pollution. O12: CairClipO3/NO2 of CAIRPOL (F)*. doi:10.2788/4277.
- Srivastava, A. K. 2003. "Detection of Volatile Organic Compounds (VOCs) Using SnO₂ Gas-Sensor Array and Artificial Neural Network." *Sensors and Actuators, B: Chemical* 96 (1–2): 24–37. doi:10.1016/S0925-4005(03)00477-5.
- Steinle, Susanne, Stefan Reis, and Clive Eric Sabel. 2013. "Quantifying Human Exposure to Air Pollution—Moving from Static Monitoring to Spatio-Temporally Resolved Personal Exposure Assessment." *Science of the Total Environment* 443. Elsevier B.V.: 184–93. doi:10.1016/j.scitotenv.2012.10.098.
- Talazac, L, J Brunet, V Battut, J.P Blanc, A Pauly, J.P Germain, S Pellier, and C Soulier. 2001. "Air Quality Evaluation by Monolithic InP-Based Resistive Sensors." *Sensors and Actuators B: Chemical* 76 (1–3): 258–64. doi:10.1016/S0925-4005(01)00579-2.
- Tham, Kwok Wai. 2016. "Indoor Air Quality and Its Effects on humans—A Review of Challenges and Developments in the Last 30 Years." *Energy and Buildings* 130. Elsevier B.V.: 637–50. doi:10.1016/j.enbuild.2016.08.071.
- Theunis, Jan, Matthias Stevens, and Dick Botteldooren. 2017. "Participatory Sensing, Opinions and Collective Awareness." In *Participatory Sensing, Opinions and Collective Awareness*, 21–46. doi:10.1007/978-3-319-25658-0.
- Tian, Fengchun, Zhifang Liang, Lei Zhang, Yan Liu, and Zhenzhen Zhao. 2016. "A Novel Pattern Mismatch Based Interference Elimination Technique in E-Nose." *Sensors and Actuators, B: Chemical* 234. Elsevier B.V.: 703–12. doi:10.1016/j.snb.2016.05.026.

- Tran, Thang Viet, Nam Trung Dang, and Wan Young Chung. 2017. "Battery-Free Smart-Sensor System for Real-Time Indoor Air Quality Monitoring." *Sensors and Actuators, B: Chemical* 248. Elsevier B.V.: 930–39. doi:10.1016/j.snb.2017.01.198.
- UNICEF. 2016. *Clear the Air for Children - The Impact of Air Pollution on Children*.
- Vergara, Alexander, Eduard Llobet, Eugenio Martinelli, Corrado Di Natale, Arnaldo D'Amico, and Xavier Correig. 2007. "Feature Extraction of Metal Oxide Gas Sensors Using Dynamic Moments." *Sensors and Actuators, B: Chemical* 122 (1): 219–26. doi:10.1016/j.snb.2006.05.028.
- Vergara, Alexander, Eugenio Martinelli, Eduard Llobet, Franco Giannini, Arnaldo D'Amico, and Corrado Di Natale. 2007. "An Alternative Global Feature Extraction of Temperature Modulated Micro-Hotplate Gas Sensors Array Using an Energy Vector Approach." *Sensors and Actuators, B: Chemical* 124 (2): 352–59. doi:10.1016/j.snb.2006.12.050.
- Vergara, Llobet, Brezmes, Ivanov, Vilanova, Gracia, Cané, and Correig. 2005. "Optimised Temperature Modulation of Metal Oxide Micro-Hotplate Gas Sensors through Multilevel Pseudo Random Sequences." *Sensors and Actuators, B: Chemical* 111–112 (SUPPL.): 271–80. doi:10.1016/j.snb.2005.06.039.
- Wang, and Brauer. 2014. "Review of Next Generation Air Monitors for Air Pollution." *School of Population and Public Health, The University of British Columbia*, 1–79. doi:10.14288/1.0132725.
- Wang, Jiayu Li, He Jing, Qiang Zhang, Jingkun Jiang, and Pratim Biswas. 2015. "Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement." *Aerosol Science and Technology* 49 (11): 1063–77. doi:10.1080/02786826.2015.1100710.
- Wetchakun, K., T. Samerjai, N. Tamaekong, C. Liewhiran, C. Siri Wong, V. Kruefu, A. Wisitsoraat, A. Tuantranont, and S. Phanichphant. 2011. "Semiconducting Metal Oxides as Sensors for Environmentally Hazardous Gases." *Sensors and Actuators, B: Chemical* 160 (1). Elsevier B.V.: 580–91. doi:10.1016/j.snb.2011.08.032.
- Williams, Ron, Amanda Kaufman, and Sam Garvey. 2015. "Next Generation Air Monitoring (NGAM) VOC Sensor Evaluation Report VOC Sensor Evaluation Report," no. May.
- Williams, Kaufman, Hanley, Rice, and Garvey. 2014. "Evaluation of Field-Deployed Low Cost PM Sensors." doi:EPA/600/R-14/464 (NTIS PB 2015-102104).
- Williams, and Kilaru. 2014. "Air Sensor Guidebook." *Environmental Protection Agency, Office of Research and Development*. doi:10.1017/CBO9781107415324.004.
- Wong, C. J., M. Z. Matjafri, K. Abdullah, H. S. Lim, and K. L. Low. 2007. "Temporal Air Quality Monitoring Using Surveillance Camera." *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2864–68. doi:10.1109/IGARSS.2007.4423441.

- Wu, Xuedong, and Zhihuan Song. 2013. "Multi-Step Prediction of Chaotic Time-Series with Intermittent Failures Based on the Generalized Nonlinear Filtering Methods." *Applied Mathematics and Computation* 219 (16). Elsevier Inc.: 8584–94. doi:10.1016/j.amc.2013.02.071.
- Xue, Yu, and Zhiqiang John Zhai. 2017. "Inverse Identification of Multiple Outdoor Pollutant Sources with a Mobile Sensor." *Building Simulation* 10 (2): 255–63. doi:10.1007/s12273-016-0322-3.
- Yamazoe, Noboru. 2005. "Toward Innovations of Gas Sensor Technology." *Sensors and Actuators, B: Chemical* 108 (1–2 SPEC. ISS.): 2–14. doi:10.1016/j.snb.2004.12.075.
- Yu, Tsang-Chu, and Chung-Chih Lin. 2015. "An Intelligent Wireless Sensing and Control System to Improve Indoor Air Quality: Monitoring, Prediction, and Preaction." *International Journal of Distributed Sensor Networks* 11 (8). Hindawi Publishing Corporation: 140978. doi:10.1155/2015/140978.
- Zampolli, S., I. Elmi, F. Ahmed, M. Passini, G. C. Cardinali, S. Nicoletti, and L. Dori. 2004. "An Electronic Nose Based on Solid State Sensor Arrays for Low-Cost Indoor Air Quality Monitoring Applications." *Sensors and Actuators, B: Chemical* 101 (1–2): 39–46. doi:10.1016/j.snb.2004.02.024.
- Zhang, Wenjie, Wei Wang, Jianhua Chen, Hongjie Liu, Tianyou Dai, Xiao-Yang Yang, Fan Zhang, Jun Lin, and Zifa Wang. 2010. "Pollution Situation and Possible Markers of Different Sources in the Ordos Region, Inner Mongolia, China." *Science of The Total Environment* 408 (3): 624–35. doi:10.1016/j.scitotenv.2009.10.033.
- Zhang, Tian, Liu, Dang, Peng, and Yin. 2013. "Chaotic Time Series Prediction of E-Nose Sensor Drift in Embedded Phase Space." *Sensors and Actuators, B: Chemical* 182. Elsevier B.V.: 71–79. doi:10.1016/j.snb.2013.03.003.
- Zhang, Yang, Simpson, Huang, Yu, Huang, Wang, et al. 2018. "Decadal Changes in Emissions of Volatile Organic Compounds (VOCs) from on-Road Vehicles with Intensified Automobile Pollution Control: Case Study in a Busy Urban Tunnel in South China." *Environmental Pollution* 233 (February). Elsevier: 806–19. doi:10.1016/J.ENVPOL.2017.10.133.

8 Appendices

8.1 Appendix 1: Commercially-available air monitoring devices

An overview of capabilities and cost of commercially-available sensor devices is provided in Table 16.

Table 16. Summary of select commercially-available air monitoring devices.

Pollutants	Range	Limit of Detection	Accuracy	Precision	Field Testing	Cost	Device
NO	0-20000 ppb	<3 ppb	±5 ppb	—	Yes	\$5500-\$9000	Geotech AQMesh
NO ₂	0-200 ppb	<5 ppb	±5 ppb	—	Yes	\$5500-\$9000	Geotech AQMesh
NO ₂	0-250 ppb	20 ppb	<30 %	±7 ppb, ±15 %	Yes	~\$60~\$5400	Cairpol CairSens
NO ₂	0.05-10 ppm	—	—	—	No	~\$500~\$650	Sensaris SensPods
NO ₂	0.05-5 ppm	—	—	—	No	\$175-\$230	Smart Citizen
NO ₂	0.05-5 ppm	—	—	—	No	\$185	Air Quality Egg
NO ₂	0-20 ppm	—	—	—	Yes	\$1000	CitiSense
NO ₂	0-5 ppm	—	—	—	No	\$199	Sensorcon Sensordrone
NO ₂	0.05-5 ppm	—	—	3 ppb	No	\$180	AirCasting Air Monitor
NO ₂	10-250 ppb	—	—	10 ppb	Yes	\$2250	Unitec SENS-IT
NO ₂	0-1ppm	—	—	± 20 ppb	Yes	\$460	Aeroqual Series 500

O ₃	0-200 ppb	<5 ppb	±5 ppb	—	Yes	\$5500-\$9000	Geotech AQMesh
O ₃	0-250 ppb	20 ppb	<30 %	±7 ppb, ±15 %	Yes	~\$60~\$5400	Cairpol CairSens
O ₃	10-1000 ppb	—	—	—	No	~\$500~\$650	Sensaris SensPods
O ₃	10-1000 ppb	—	—	—	No	\$185	Air Quality Egg
O ₃	0-1 ppm	—	—	—	Yes	\$1000	CitiSense
O ₃	2 ppb - 10 ppm	—	—	2 ppb	Yes	\$4900	2B Technologies
O ₃	0 -500ppb	—	—	± 5.0 ppb	Yes	\$460	Aeroqual Series 500
CO	100-8000 ppb	—	—	200 ppb	Yes	\$2250	Unitec SENS-IT
CO	0-50000 ppb	<5 ppb	±10 ppb	—	Yes	\$5500-\$9000	Geotech AQMesh
CO	0-20 ppm	0.05 ppm	<25 %	±0.05 ppm, ±15 %	Yes	~\$60~\$5400	Cairpol CairSens
CO	1-1000 ppm	—	—	—	No	~\$500~\$650	Sensaris SensPods
CO	1-1000 ppm	—	—	—	No	\$175-\$230	Smart Citizen
CO	1-1000 ppm	—	—	—	No	\$185	Air Quality Egg
CO	0-2000 ppm	—	—	—	Yes	\$1000	CitiSense
CO	0-1000 ppm	—	—	—	No	\$199	Sensorcon Sensordrone

CO	30~1000 ppm	—	—	—	No	\$180	AirCasting Air Monitor
CO	0 - 250 ppb	—	—	± 50 ppb	Yes	\$460	Aeroqual Series 500
CO ₂	0-5000 ppm	—	±30 ppm	±5 %	No	~\$500~\$650	Sensaris SensPods
CO ₂	0-2000 ppm	—	±50 ppm	—	No	~\$830	Envirologger
CH ₂ O	0-1000 ppb	10 ppb	<30 %	±5 ppb, ±20 %	Yes	~\$60~\$5400	Cairpol CairSens
VOCs	0-16 ppm	10 ppb	<30 %	±10 ppb, ±15 %	Yes	~\$60~\$5400	Cairpol CairSens
VOCs	1-1000 ppm	—	—	—	No	\$185	Air Quality Egg
PM _{2.5}	0-250 µg/m ³	5 µg/m ³	±50 %	±5 µg/m ³ , ±10 %	Yes	~\$60~\$5400	Cairpol CairSens
PM ≥1 µm	0-28,000 pcs/L	—	—	—	No	~\$500~\$650	Sensaris SensPods
PM ≥1 µm	0-28,000 pcs/L	—	—	—	No	\$185	Air Quality Egg
PM _{2.5}	—	—	—	—	No	\$99	Speck / GPSpeck

8.2 Appendix 2: Design for the AirSENCE sensor device

8.2.1 Electrical design for the AirSENCE sensor device

This electrical system of each AirSENCE device (Rev1-2.5) was comprised of three system level modular components: 1) data logging unit (composed of a microcontroller board, a data logging board and

connection board); 2) power interface board, and 3) two sensor boards (Figure 74). The modular, component-based design was used to simplify device troubleshooting and maintenance.

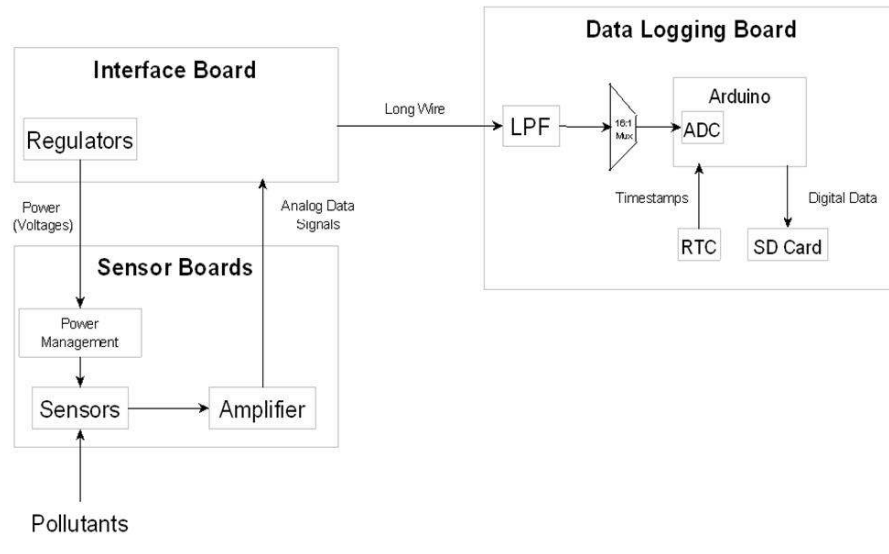


Figure 74. System Level Diagram of Electrical Components.

The data logging unit consisted of three separate printed circuit boards (Figure 75, Figure 76): 1) a microcontroller board (i.e. commercial Arduino Uno board), 2) a data logging board and 3) a connection board containing a multiplexer, low pass filter and cable connectors for linking to the interface board. Low pass filter was used to reduce the noise in sensor signals. The measurements were passed through a 16:1 analog multiplexer, allowing for probing of 16 sensors with 4 I/O pins. An analog to Digital Converter (ADC) on the microcontroller board was used to convert analog sensor measurements to digital format. The data then gets stored on a micro-SD card using a micro-SD read/write breakout board. A Real Time Clock (RTC) chip was used to generate time stamps and store them along with the sensor data. During PCB design, thin traces were used for signal and thicker traces - for power lines. Signal traces were routed to avoid thermal noise from the heat generated by chips.

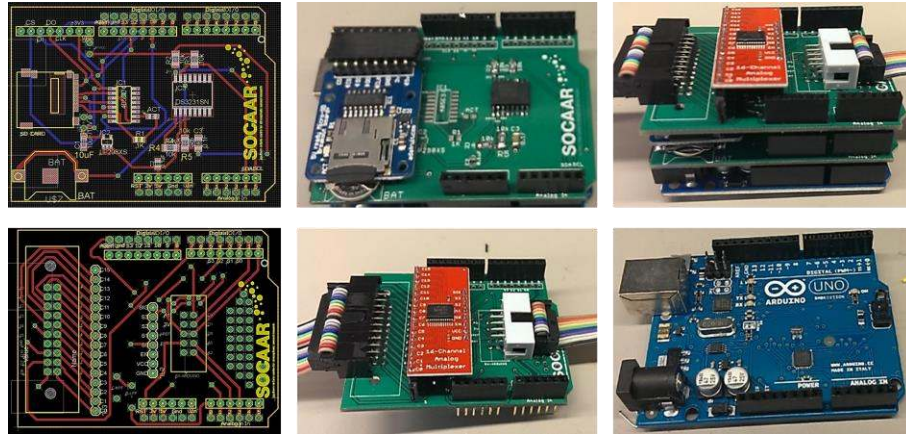


Figure 75. The design for circuit boards comprising data logging unit as well as images of completed and assembled boards.

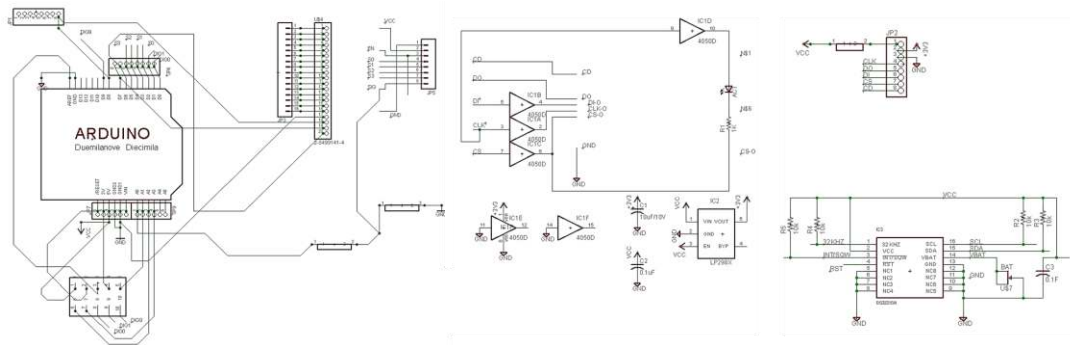


Figure 76. Data logging unit circuit schematics.

The interface board (Figure 77, Figure 78) has two key roles:

- 1) To provide the sensors with power at the required voltage and current levels;
- 2) To route the signals from the two sensor boards to the data logging board.

While most sensors required the same voltage as given by AC adaptor (5V), some needed voltages lower or higher. Switching converters were used as power regulators.

Finally, a footprint for an additional microcontroller is available on the interface board and was included for implementing power management functionality in the device.

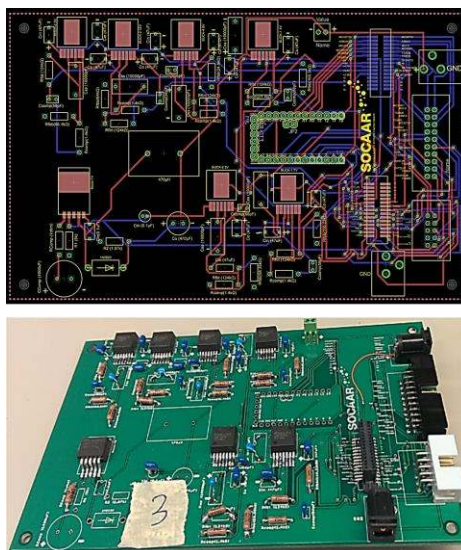


Figure 77. The design for the power interface board as well as image of a completed and assembled board.

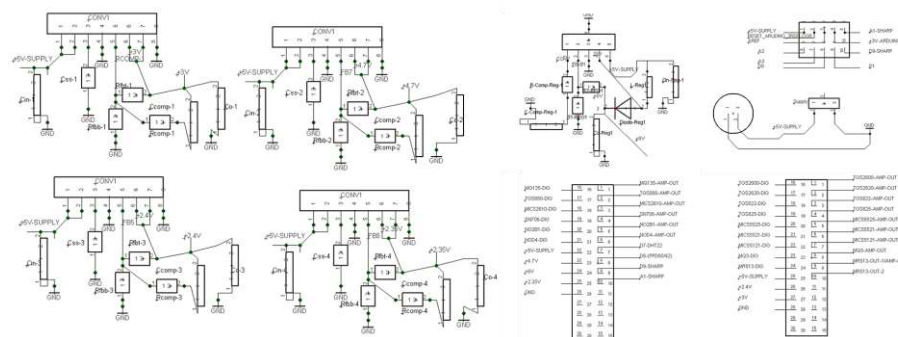


Figure 78. Power interface board circuit schematics.

Each AirSENCE device was composed of two sensor boards, each with different types of sensors (Figure 79, Figure 80). The sensors on boards were grouped based on two factors: 1) Power consumption and 2) Types of pollutants monitored. Metal oxide semiconductor gas sensors consume significant amount of power (500-800 mW) due to the heater integrated in the sensor. The heater catalyzed the redox reactions happening on sensor surface. As a result, it was important to balance the current draw of the two sensor boards and to ensure similar heat gradient across the two sides of the device. Since power consumption of sensors was correlated to the amount of heat they generate, balancing power consumption also ensured balanced convection on the two sides of the device. In addition, the sensors were grouped according to the pollutants they measured to ensure consistent readings. Provision for breakout boards for power management and signal amplification were also made.

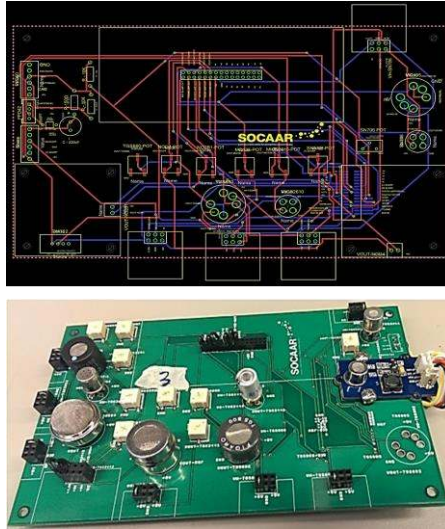


Figure 79. The design for one of the sensor boards as well as image of a completed and assembled board.

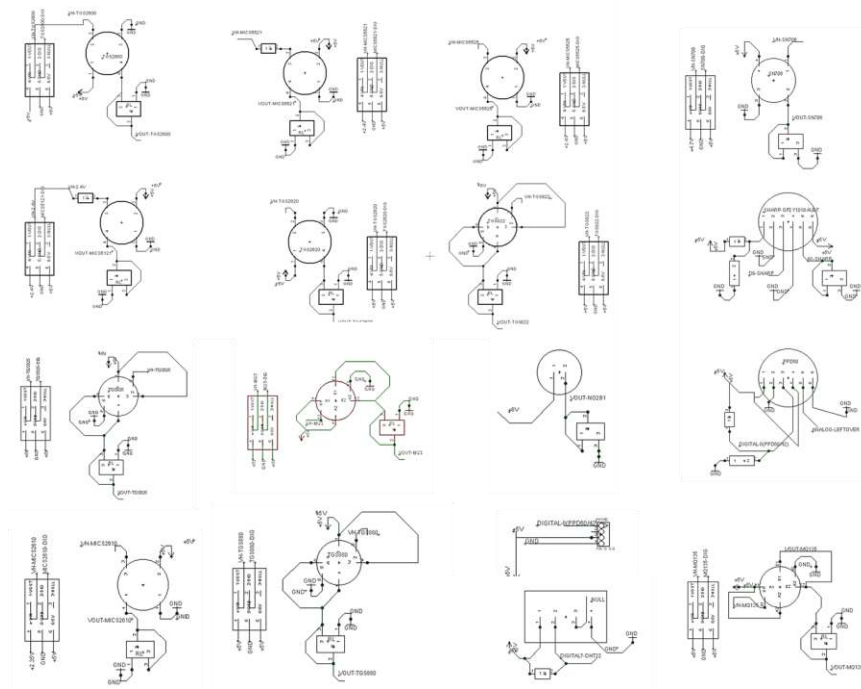


Figure 80. Sensor board circuit schematics.

A complete list of evaluated sensors is provided in Table 17 and the list of other components used in design is provided in Table 18.

Table 17. Evaluated sensor list.¹²

Manufacturer	Model	Operation	Price (\$CAN) for 1-9 units	Target Analyte
Aosong Electronics	DHT22	Thermistor, capacitive sensor	\$15	Temperature, humidity
Figaro	TGS2602	MOS	\$15	Methane, CO, iso-butane, ethanol, hydrogen
Figaro	TGS2600	MOS	\$15	Methane, CO, iso-butane, ethanol, hydrogen
Figaro	TGS2620	MOS	\$15	Methane, CO, iso-butane, ethanol, hydrogen
Figaro	TGS822	MOS	\$15	Organic solvents, acetone, methane, CO, isobutane
Figaro	TGS825	MOS	\$57	Hydrogen sulfide
Figaro	TGS800	MOS	\$15	CO, methane, isobutane, hydrogen and ethanol
Figaro	TGS823	MOS	\$24	Acetone, methane, CO, iso-butane, ethanol, hydrogen
Figaro	TGS882	MOS	\$15	Fumes from food (alcohol, odor)
Figaro	TGS2201	MOS	\$15	Gasoline and Diesel Exhaust Gas Sensor
Figaro	TGS2611	MOS	\$15	Methane, iso-butane, ethanol, hydrogen
Figaro	TGS880	MOS	\$15	CO, methane, isobutane, ethanol, cooking vapours
Figaro	TGS826	MOS	\$15	Ammonia, ethanol
Figaro	TGS832	MOS	\$25	Chlorofluorocarbons
Figaro	TGS2442	MOS	\$15	CO
E2V / MICS	MICS5525	MOS	\$10	CO
E2V / MICS	MICS5521	MOS	\$10	CO, hydrocarbons, volatile organics
E2V / MICS	MICS5121	MOS	\$10	CO, hydrocarbons, volatile organics (protective filter)
E2V / MICS	MICS2610	MOS	\$10	O ₃

¹² <https://www.yumpu.com/en/document/view/379764/price-delivery-information-for-figaro-gas-sensors>

E2V / MICS	MICS2710	MOS	\$10	NOx
E2V / MICS	MICS2611	MOS	\$10	O ₃
Synkera	P/N706	MOS	\$30	NOx
Synkera	P/N707	MOS	\$30	Volatile organics (methanol, dicholoromethane, hexane, ethyl acetate, isopropyl acetate)
Hanwei	MR513	MOS	\$23	Alcohol
Hanwei	MQ3	MOS	\$5	Alcohol, benzene, CH ₄ , hexane, LPG, CO
Hanwei	MQ135	MOS	\$8	NH ₃ , NOx, alcohol, benzene, smoke, CO ₂ , CO
Hanwei	MQ131	MOS	\$20	O ₃
Hanwei	MQ7	MOS	\$9	CO
Applied sensor	ASMLN	MOS	\$30	NOx
Sharp	GP2Y1010 AU0F	Optical	\$12	Particulate matter
Shinyei	PPD 42	Optical	\$12	Particulate matter
Grove	SEN 21723	N-MOS	\$15	Sensor

Table 18. Bill of materials for components used in design.

Part	What	Value	Model	Amount	Price (\$CAN)
Power	Jack		PJ-102B	1	1.17
	Switch			10	5.84
	Indication LED	Red	WP3A10HD	10	0.97
	LED Resistor	1k		50	1.27
SMPS	IC		RT7273	2	3.84
	Capacitor	10uF	CL21A106KOQNNNE	50	4.03
		22uF	CL21A226MQQNNNE	50	5.5
		3.3uF	CL21A335KPFNNNE	25	3.44
	Resistor	100k	ERJ-6ENF1003V	50	1.44
		383k	ERJ-6ENF3833V	50	1.44

5V	Capacitor	100nF	CL21B104KACNNNC	50	1.2
		4.7nF	CL21B472KBANNNC	50	1.27
		5.6nF	CL21B562KBANFNC	50	1.4
		22uF		-	-
		470pF	CL21B471KCANNNC	50	0.8
	Resistor	24k	ERJ-6GEYJ243V	50	0.98
		43k	ERJ-6GEYJ433V	50	0.98
		40.2k	ERJ-6ENF4022V	50	1.44
		7.68k	ERJ-6ENF7681V	50	1.44
	Inductor	6.8uH	SRR1240-6R8M	10	8.72
8V	Capacitor	4.7nF		-	-
		5.6nF		-	-
		100nF		-	-
		470pF		-	-
		22uF		-	-
	Resistor	24k		-	-
		51k	ERJ-6GEYJ513V	50	0.98
		40.2k		-	-
		4.42k	ERJ-6ENF4421V	50	1.44
	Inductor	6.8uH		-	-
4.7V	Capacitor	4.7nF		-	-
		5.6nF		-	-
		100nF		-	-
		470pF		-	-
		22uF		-	-
	Resistor	24k		-	-
		51k		-	-
		40.2k		-	-
		8.25k		50	1.44
	Inductor	6.8uH		-	-
2.4V	Capacitor	4.7nF		-	-
		100nF		-	-

		470pF		-	-
		22uF		-	-
	Resistor	20k	ERJ-6ENF2002V	50	1.44
		51k		-	-
		40.2k		-	-
		20k		-	-
	Inductor	4.7uH	SRN5020-4R7M	10	3
2.35V	Resistor	21k	ERJ-6ENF2102V	50	1.44
	Sensor Board Connector		A33573-ND	2	14.36
	LED/Switch Connector		1984617	10	3.8

8.2.2 Code for AirSENCE sensor device

8.2.2.1 Code for logging sensor data

```
#include <Wire.h> // Wire library that helps the Arduino with i2c
#include <Streaming.h> // Streaming C++-style Output with Operator << from
http://arduinoiana.org/libraries/streaming/

//-----
// CONFIGURATION
//-----

// define FAST sensor sampling interval
unsigned long INTERVAL_FAST = 5000; // sample duration (ms)30s, david holstius used 60000
unsigned long last_time_fast_ms; // time of last check/output

char CSV_HEADER[] = "Timestamp,Sensor,Variable,Value";

//-----
// COMPONENT SETUP
//-----
```

```
//----- MUX

int MUX_BASE = 0; // analog pin to interact with mux
int MUX_CONTROL[] = {2, 3, 4, 5}; // digital pins for mux pin selection
int mux(int channel) {
    // select channel on multiplexer
    for (int i = 0; i < 4; i++) { digitalWrite(MUX_CONTROL[i], bitRead(channel, i)); }
    // return base pin for analogRead/Write to use
    return(MUX_BASE);
}

void setup_mux() {
    // set digital pins controlling multiplexer to OUTPUT
    for (int i = 0; i < 4; i++) {
        pinMode(MUX_CONTROL[i], OUTPUT);
    }
}

//----- SD card (logging)

#include "SD.h" //SD library to talk to the card
#define SD_PIN 10 // for the data logging shield, we use digital pin 10 for the SD cs line
#define SD_INTERVAL_SYNC 1000 // mills between calls to flush() - to write data to the card
uint32_t sd_last_time = 0; // time of last sync()
File logfile; // the logging file

// error function (ex. couldn't write to the SD card or open it)
void error_sd(char *msg) {
    log_error("SDcard", msg);
    while(1); // sits in a while loop forever to halt setup
}

// setup
```

```

void setup_sd() {

    // initialize the SD card and find a FAT16/FAT32 partition
    log_info("SDcard", "initializing...");

    pinMode(SD_PIN, OUTPUT); // default chip select pin is set to output

    if (!SD.begin(SD_PIN)) { // see if the card is present and can be initialized
        error_sd("card failed or not present");
        return; // don't do anything else
    }

    log_info("SDcard", "initialized");

    // create a new file, LOGGERnnnn.csv (nnnn is a number), make a new file every time the Arduino
    starts up
    char filename[] = "0000.csv";
    for (uint16_t i = 0; i < 10000; i++) {

        for (int j = 0; j < 4; j++) {
            filename[4-j-1] = ((int)(i / pow(10, j)) % 10) + '0';
        }

        if(!SD.exists(filename)) { // only open a new file if it doesn't exist
            logfile = SD.open(filename, FILE_WRITE); // Unix style command flags - logfile.open() procedure
            FILE_WRITE - create the file and write data to it
            break; // leave the loop!
        }
    }

    if (! logfile) {
        error_sd("could not create file");
    }
}

```

```

logfile << CSV_HEADER << endl;
log_info("SDcard", filename);

sd_last_time = millis();
}

//----- Chronodot (timekeeping)

#include "Chronodot.h" // Chronodot RTC library
Chronodot RTC;

// setup
void setup_rtc() {
    // kick off the RTC initializing the Wire library - poking the RTC to see if its alive
    Wire.begin();
    RTC.begin(); // Chronodot

    if (!RTC.begin()) {
        logfile.println("RTC failed");
        Serial.println("RTC failed");
    }
}

// log
void log_time() {
    DateTime now = RTC.now();
    int year = now.year();
    int month = now.month();
    int day = now.day();
    int hours = now.hour();
    int mins = now.minute();
    int secs = now.second();

```

```

Serial << "\" << year;
Serial << "-" << ((month<10)?"0:") << month;
Serial << "-" << ((day<10)?"0:") << day;
Serial << " " << ((hours<10)?"0:") << hours;
Serial << ":" << ((mins<10)?"0:") << mins;
Serial << ":" << ((secs<10)?"0:") << secs;
Serial << "\", ";

logfile << "\" << year;
logfile << "-" << ((month<10)?"0:") << month;
logfile << "-" << ((day<10)?"0:") << day;
logfile << " " << ((hours<10)?"0:") << hours;
logfile << ":" << ((mins<10)?"0:") << mins;
logfile << ":" << ((secs<10)?"0:") << secs;
logfile << "\", ";
}

//----- DHT22 (temperature)

#include "DHT.h"
#define DHT_PIN 7
#define DHT_TYPE DHT22 // DHT 22 (AM2302)
DHT dht(DHT_PIN, DHT_TYPE);

// setup
void setup_dht() {
  dht.begin();
}

// log
void log_dht() {
  // Reading temperature or humidity takes about 250 milliseconds!
  // Sensor readings may also be up to 2 seconds 'old' (its a very slow sensor)
  float h = dht.readHumidity();

```

```

float t = dht.readTemperature();
// check if returns are valid, if they are NaN (not a number) then something went wrong!
if (isnan(t) || isnan(h)) {
    log_error("DHT22", "failed to read");
} else {
    log_measurement("DHT22", "humidity", h);
    log_measurement("DHT22", "temperature", t);
}
}

//----- Sinyei (particles)

#define SNY_PIN 8
unsigned long sny_duration;
unsigned long sny_lowocc_us = 0; // sum of time spent LOW

unsigned long SNY_INTERVAL_ms = 30000; // sample duration (ms)30s, david holstius used 60000
unsigned long sny_last_time;      // time of last serial dump

// setup
void setup_sinyei() {
    pinMode(SNY_PIN, INPUT);
    sny_last_time = millis();
    sny_lowocc_us = 0;
}

// reading
void log_sinyei(unsigned long lowocc_us, unsigned long sampletime_ms) {
    float ratio; // fraction of time spent LOW
    float concentration = 0;

    ratio = lowocc_us / (sampletime_ms * 1000.0) * 100.0;
    concentration = 1.1 * pow(ratio, 3) - 3.8 * pow(ratio, 2) + 520 * ratio + 0.62; // using spec sheet curve

```

```

log_measurement("Sinyei", "samptime_ms", samptime_ms);
log_measurement("Sinyei", "low_occupancy_duration_us", lowocc_us);
log_measurement("Sinyei", "occupancy_ratio", ratio);
log_measurement("Sinyei", "concentration", concentration);
}

//----- Sharp (particles)

// connected to analog 1 and controlled by digital 9
#define SRP_PIN_DUST 1
#define SRP_PIN_LED 9 // led Power is any digital pin on the arduino connected to Pin 3 on the sensor
(can be 2, 4,...)

int srp_delay1_us = 280; // delays are in microseconds
int srp_delay2_us = 40;
int srp_delayoff_us = 9680;

// setup
void setup_sharp() {
  pinMode(SRP_PIN_LED, OUTPUT);
}

// log
void log_sharp() {
  digitalWrite(SRP_PIN_LED, LOW); // power on the LED
  delayMicroseconds(srp_delay1_us);
  int dustVal = analogRead(SRP_PIN_DUST); // read the dust value
  delayMicroseconds(srp_delay2_us);
  digitalWrite(SRP_PIN_LED, HIGH); // turn the LED off
  delayMicroseconds(srp_delayoff_us);

  log_measurement("Sharp", "dust", dustVal);

```

```

}

//----- Arduino voltage supply

// read Arduino voltage
long read_vcc() {
    // Read 1.1V reference against AVcc
    // set the reference to Vcc and the measurement to the internal 1.1V reference
    #if defined(__AVR_ATmega32U4__) || defined(__AVR_ATmega1280__) ||
defined(__AVR_ATmega2560__)
        ADMUX = _BV(REFS0) | _BV(MUX4) | _BV(MUX3) | _BV(MUX2) | _BV(MUX1);
    #elif defined (__AVR_ATtiny24__) || defined(__AVR_ATtiny44__) || defined(__AVR_ATtiny84__)
        ADMUX = _BV(MUX5) | _BV(MUX0);
    #elif defined (__AVR_ATtiny25__) || defined(__AVR_ATtiny45__) || defined(__AVR_ATtiny85__)
        ADMUX = _BV(MUX3) | _BV(MUX2);
    #else
        ADMUX = _BV(REFS0) | _BV(MUX3) | _BV(MUX2) | _BV(MUX1);
    #endif
    delay(2); // Wait for Vref to settle
    ADCSRA |= _BV(ADSC); // Start conversion
    while (bit_is_set(ADCSRA,ADSC)); // measuring
    uint8_t low = ADCL; // must read ADCL first - it then locks ADCH
    uint8_t high = ADCH; // unlocks both
    long result = (high<<8) | low;
    result = 1125300L / result; // Calculate Vcc (in mV); 1125300 = 1.1*1023*1000
    return result; // Vcc in millivolts
}

// log
void log_vcc() {
    float vcc = read_vcc();
    log_measurement("Arduino", "voltage", vcc);
}

```



```

}

//-----
// LOGGING
//-----

void log_measurement(String sensor, String variable, float value) {
  log_time();
  Serial << sensor << "," << variable << ",";
  Serial.print(value, 4);
  Serial << endl;
  logfile << sensor << "," << variable << ",";
  logfile.print(value, 4);
  logfile << endl;
}

void log_error(String sensor, String msg) {
  log_status(sensor, "ERROR", msg);
}

void log_info(String sensor, String msg) {
  log_status(sensor, "INFO", msg);
}

void log_status(String sensor, String category, String msg) {
  log_time();
  Serial << sensor << "," << category << "," << msg << endl;
  logfile << sensor << "," << category << "," << msg << endl;
}

//-----
// SETUP
//-----

```

```

void setup(void) {

    // setup serial
    Serial.begin(9600);
    Serial << CSV_HEADER << endl;

    // setup RTC
    setup_rtc(); // all logging depends on RTC, so start it up first

    // setup SD card
    setup_sd();

    // setup MUX
    setup_mux();

    // setup DHT
    setup_dht();

    // setup Sinyei
    setup_sinyei();

    // setup Sharp
    setup_sharp();

    // set fast sensor clock
    last_time_fast_ms = millis(); // get the current time

}

//-----
// MAIN LOOP
//-----

```

```

void loop() {

//main loop: 1)Wait until time for the next reading, 2)Ask for the current time and date from RTC,
//3)Log the time and date to SD, 4)Read sensors, 5)Log readings to the SD, 6)Sync data to the card at
right time

sny_duration = pulseIn(SNY_PIN, LOW);
sny_lowocc_us += sny_duration;

//----- read/log Sinyei sensor (30s interval)

if ((millis() - sny_last_time) > SNY_INTERVAL_ms) {

// log Sinyei
log_sinyei(sny_lowocc_us, millis() - sny_last_time);

// reset
sny_lowocc_us = 0;
sny_last_time = millis();

}

//----- read/log other sensors (fast 10s interval)

if ((millis() - last_time_fast_ms) > INTERVAL_FAST) {

// log Arduino voltage
log_vcc();

// log DHT

```

```

log_dht();

// log Sharp
log_sharp();

// log MUX
int num_sensors = 16;
int sensor_value;
String sensor_name = "channel_";

for (int channel = 0; channel < num_sensors; channel++) {
    sensor_value = analogRead(mux(channel));
    log_measurement("Mux", sensor_name + ((channel<10)?"0":"") + channel, sensor_value);
}

// reset clock to current time
last_time_fast_ms = millis();

}

//----- sync SD card

// write data to disk! Don't sync too often (2048 bytes of I/O SD card, a lot of power, takes time)
if ((millis() - sd_last_time) > SD_INTERVAL_SYNC) {
    logfile.flush();
    sd_last_time = millis();
}
}

```

8.2.2.2 Code for data merging and cleaning

```

# Aggregate sensor array time series data from multiple devices stored
# as flat CSV files with columns [Port,] Timestamp, Sensor, Variable, Value

```

```
#-----
```

```
from __future__ import division
```

```
import os
```

```
import sys
```

```
import csv
```

```
import glob
```

```
from datetime import datetime, timedelta, time
```

```
from collections import defaultdict
```

```
import logging
```

```
logger = logging.getLogger()
```

```
# default fieldnames
```

```
FIELDNAMES_IN = ['Deployment', 'Device', 'Location', 'Port', 'Timestamp', 'Sensor', 'Variable', 'Value']
```

```
FIELDNAMES_OUT = ['Timeslot', 'Deployment', 'Device', 'Location', 'Port', 'NumObservations', 'Sensor',  
'Variable', 'Value']
```

```
FIELDNAMES_IN_REQUIRED = FIELDNAMES_IN[-4:]
```

```
MAX_FIELDS = 5
```

```
PARSERS = {
```

```
    'Timestamp': lambda v: datetime.strptime(v, '%Y-%m-%d %H:%M:%S'),
```

```
    'Value': lambda v: float(v),
```

```
}
```

```
DEFAULTS = {
```

```
    'Port': None,
```

```
}
```

```
def get_timeslot(timestamp, interval):
```

```
    if not isinstance(interval, timedelta):
```

```
        raise TypeError('Interval must be a timedelta object')
```

```

if interval > timedelta(days=1):
    raise ValueError('Intervals longer than a day are not supported')

if interval < timedelta(seconds=1):
    raise ValueError('Intervals shorter than a second are not supported')

if interval.total_seconds() < 0:
    raise ValueError('Negative intervals are not supported')

if (timedelta(days=1).total_seconds() % interval.total_seconds()) != 0:
    raise ValueError('The chosen interval must evenly divide a day')

# Construct baseline at the start of the day,
# with all attributes potentially affected by the interval set to zero
baseline = datetime(year=timestamp.year, month=timestamp.month, day=timestamp.day)

# Find and center timeslot
slots_elapsed = int((timestamp - baseline).total_seconds() // interval.total_seconds())
timeslot = baseline + interval * slots_elapsed # Start
timeslot += interval // 2 # Center

return timeslot

def get_meta(path):
    levels = []
    path = os.path.abspath(path)
    for i in range(3):
        path = os.path.dirname(path)
        levels.append(os.path.basename(path))
    deployment, location, device = reversed(levels)
    return {'Deployment': deployment, 'Location': location, 'Device': device}

```

```

def csv_clean(iterable):
    for i, line in enumerate(iterable):
        line = line.replace('\x00', '')
        # Construct default CSV header if missing
        if i == 0 and 'Timestamp' not in line:
            logger.debug('Using default header')
            fields = csv.reader([line]).next()
            if len(fields) > MAX_FIELDS:
                logger.error("Too many fields ({n}) to be a valid input CSV file".format(n=len(fields)))
                break
            yield ', '.join(FIELDNAMES_IN[-len(fields):])
        else:
            yield line

```

```

def csv_read(path):
    metadata = get_meta(path)
    logger.info('Deployment: "{Deployment}"', Location: "{Location}", Device:
"{Device}"").format(**metadata))
    with open(path, "rb") as handle:
        reader = csv.DictReader(csv_clean(handle))
        while True:
            # Attempt to read next row
            try:
                csv_row = reader.next()
                i = reader.line_num
            except StopIteration:
                break
            except Exception as e:
                logger.warning('Row {i}: Failed to parse row: "{msg}"'.format(i=reader.line_num,
msg=e.message))
                continue

```

```

# Pre-process row
row = {}
row.update(DEFAULTS)
success = True
for k, v in csv_row.items():
    try:
        row[k.strip()] = v.strip().strip('"')
    except Exception as e:
        logger.warning('Row {i}: Invalid data {k}="{v}"; skipping'.format(i=i, k=k, v=v))
        success = False
        break
if not success:
    continue # Skip

# Check for required fields
if not all(k in row for k in FIELDNAMES_IN_REQUIRED):
    logger.error(
        'Row {i}: Missing fields: {fields}'.format(
            i=i, fields=', '.join('{k}'.format(k=k) for k in FIELDNAMES_IN_REQUIRED if k not in row)
        )
    )
    break # Abandon ship!

# Skip diagnostic data
if row['Variable'] in ['INFO', 'ERROR', 'Variable']:
    logger.debug(
        'Row {i}: Skipping diagnostic data {k}="{v}"'.format(i=i, k=row['Variable'], v=row['Value'])
    )
    continue

# Parse values
try:

```



```

        for k, v in row.items():
            row[k] = PARSERS.get(k, lambda _: _)(v)
    except Exception as e:
        logger.warning(
            'Row {i}: Could not parse {k}="{v}" ({msg}); skipping'.format(i=i, k=k, v=v, msg=e.message)
        )
        continue

    # Override metadata
    row.update(metadata)

    yield row


def csv_aggregate(rows, interval):
    block = None
    timeslot_current = None
    acq_keys = 'Deployment', 'Device', 'Location', 'Port', 'Sensor', 'Variable'
    for row in rows:
        timestamp = row['Timestamp']
        timeslot = get_timeslot(timestamp, interval)

        if timeslot != timeslot_current:
            if block is not None:
                for acq_values, values in sorted(block.items()):
                    _row = dict(zip(acq_keys, acq_values))
                    _row.update({
                        'Timeslot': timeslot,
                        'NumObservations': len(values),
                        'Value': sum(values) / len(values)
                    })
                yield _row
            block = defaultdict(list)

```

```

        timeslot_current = timeslot

    block[tuple(row[k] for k in acq_keys)].append(row['Value'])

def csv_merge(src_paths, dst_path, interval=timedelta(seconds=15), min_size=1024): # min_size in KiB
    with open(dst_path, 'wb') as handle:
        writer = csv.DictWriter(handle, fieldnames=FIELDNAMES_OUT)
        writer.writeheader()
        for src_path in src_paths:
            if os.stat(src_path).st_size < min_size*1024:
                continue
            logger.info('Processing "{path}"'.format(path=src_path))
            for row in csv_aggregate(csv_read(src_path), interval):
                writer.writerow(row)

def run_aggregate(sources, filename, interval, units, min_size):
    filename = filename.format(interval=interval, units=units)

    formatter = logging.Formatter("%(asctime)15s - %(levelname)s - %(message)s")
    fh = logging.FileHandler(filename=filename+'.log', mode='w')
    fh.setLevel(logging.INFO)
    fh.setFormatter(formatter)
    ch = logging.StreamHandler()
    ch.setLevel(logging.DEBUG)
    ch.setFormatter(formatter)
    logger.setLevel(logging.DEBUG)
    logger.addHandler(fh)
    logger.addHandler(ch)

    sources = sum((glob.glob(source) if '*' in source else [source] for source in sources), [])
    logger.debug('Starting to aggregate data')

```

```

csv_merge(sorted(sources), filename, interval=timedelta(**{units: interval}), min_size=min_size)
logger.debug('Done')

def main(argv=sys.argv[1:]):
    import argparse

    parser = argparse.ArgumentParser(description='Merge time series data from CSV files.')
    parser.add_argument('-o', '--output', help='output filename', type=str, default='merged-
{interval}_{units}.csv')
    parser.add_argument('-i', '--interval', metavar='interval', type=int, help='time window to average
values over (in the specified unit)', required=True)
    parser.add_argument('-u', '--units', metavar='units', choices=('seconds', 'minutes', 'hours'), help='unit
to use for interval (seconds, minutes, or hours)', required=True)
    parser.add_argument('-m', '--min-size', metavar='size', type=int, help='minimum file size to process (in
KiB)', default=1024)
    parser.add_argument('sources', metavar='file', nargs='+', help='files to process')
    args = parser.parse_args(argv)
    run_aggregate(sources=args.sources, filename=args.output, interval=args.interval, units=args.units,
min_size=args.min_size)

if __name__ == '__main__':
    main()

```

8.2.3 Mechanical system design for the AirSENCE sensor device

The key considerations for the AirSENCE mechanical design included: 1) Ensuring consistent and unobstructed airflow through the device, 2) reducing particle sample losses and reactions on the device surface, and 3) providing a rugged and weatherproof design for device deployment in field.

The AirSENCE device consisted of two inlets at the bottom of the device and two outlets at the top of the device. The inlets were positioned at the bottom section of the device to take advantage of

convection as an additional mechanism for air circulation through the device. Two fans at the top of the device were used to sample the air by generating negative pressure within the device and thereby drawing the air through the inlets. The negative pressure sampling was used to avoid particle sample losses through impaction on fan blades. The particle sensors were also positioned close to inlets to ensure reliable sample.

The inlets and fans were positioned along the two sides of the device such that the laminar airflow was directed in two channels along the two sensor boards, while minimizing the contact with other circuits. Since humidity and salts in the air can cause corrosion, it was important to reduce the contact between the sensitive circuits and ambient air. Furthermore, it was important to ensure sufficient clearance for unobstructed airflow through pollution sensors, while keeping the design compact and reducing dead air pockets to ensure quick response time. The inlets were positioned perpendicular to the outlets to enable the device to be mounted on windowsill while providing for the vertical airflow.

Modular design was used for the enclosure and mounting fixtures to enable easy maintenance while ensuring rugged enclosure (Figure 81). The case consisted of 8 aluminum parts joined together by screws. The device was composed of two internal compartments: a sensor compartment and a microcontroller compartment. Two bent aluminum sheets were used as device walls and a top panel was used to protect the device from the rain. The top panel also contained a hook to enable mounting from a tripod.

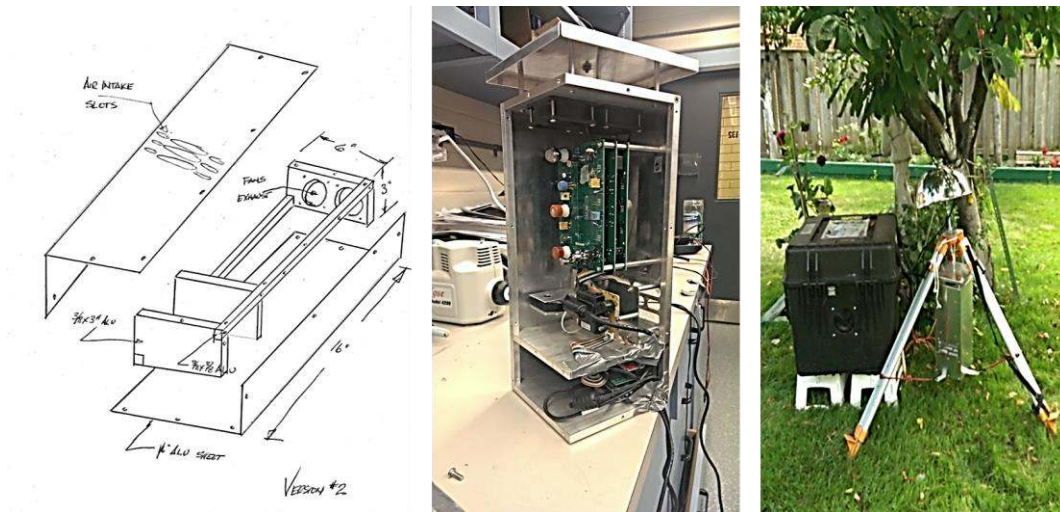


Figure 81. A sketch of mechanical design as well as implemented device and image of deployment at a site.

Two small and lightweight versions of the container were also designed to accommodate the original sensor boards, making them interchangeable between stationary and portable versions of the device (Figure 82). The airflow was generated by a fan as in the original design. The design included a portable 11Ah battery pack that ensured that the device could be powered continuously for six hours. The dimensions of the new portable versions of the device were 38x15x15cm (device version 1) and 27x19x9cm (device version 2) compared to the original device dimensions of 45x19x19cm.

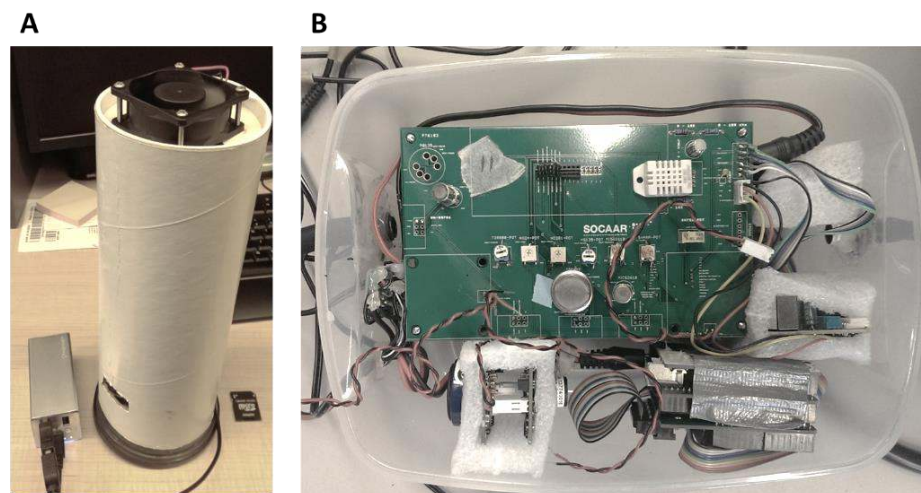


Figure 82. Two small versions of the device enclosure were designed and built. Version 1 was designed to be included in backpack used to monitor the exposure of runners. Version 2 contains a bigger range of sensors and was designed to fit into the lung transplant patient exposure assessment kit.

8.3 Appendix 3: Finalized calibration model predictor subsets and coefficients

Sample calibration coefficients for selected air pollutant regression models are provided in Table 19.

Table 19. Calibration model coefficients for each of the target pollutants for a typical AirSENCE device. A mode of five AirSENCE devices is shown for each predictor. Data source: Fall 2015, AirSENCE Rev2.

	CO	CO ₂	NO _x	O ₃	PM _{2.5}
Temperature	-9.2	-2.1	-1.3	1.3	
Humidity	-2.4	-0.6	-0.4	0.0	
TGS2600					-7.7
TGS2620	275.4				31.4
TGS822	695.8		58.9		24.8
TGS880					19.9
P/N706	273.9	116.4	48.9	-88.0	-26.5
GP2Y1010AU0F					11.1
PPD42					-0.4
Intercept	278.3	481.8	30.9	23.2	-6.9

8.4 Appendix 4: Precision and accuracy analysis

An overview of pollutant model performance with respect to accuracy and precision is provided in Figure 83, Figure 84, Figure 85 and Figure 86.

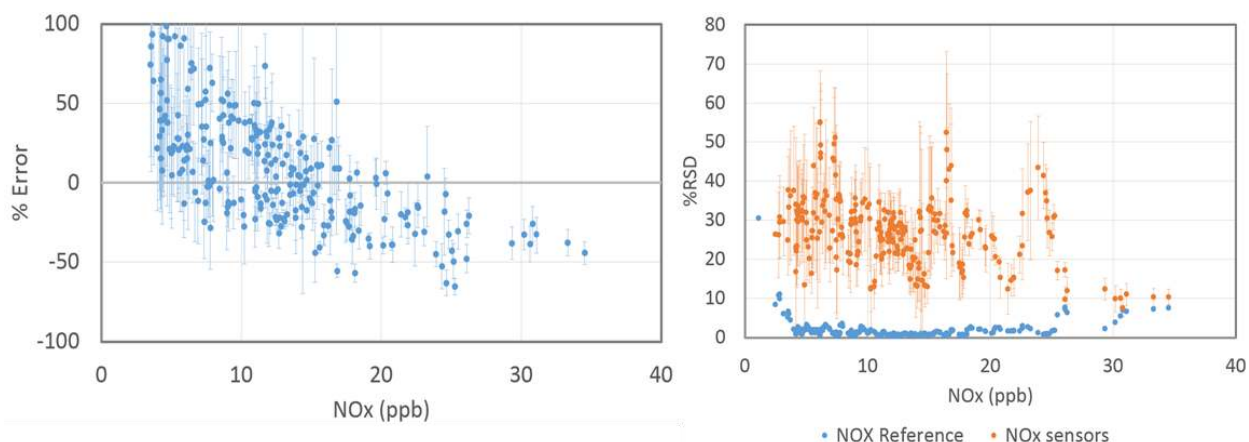


Figure 83. Evaluation of accuracy and precision of NO_x model with change in NO_x concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

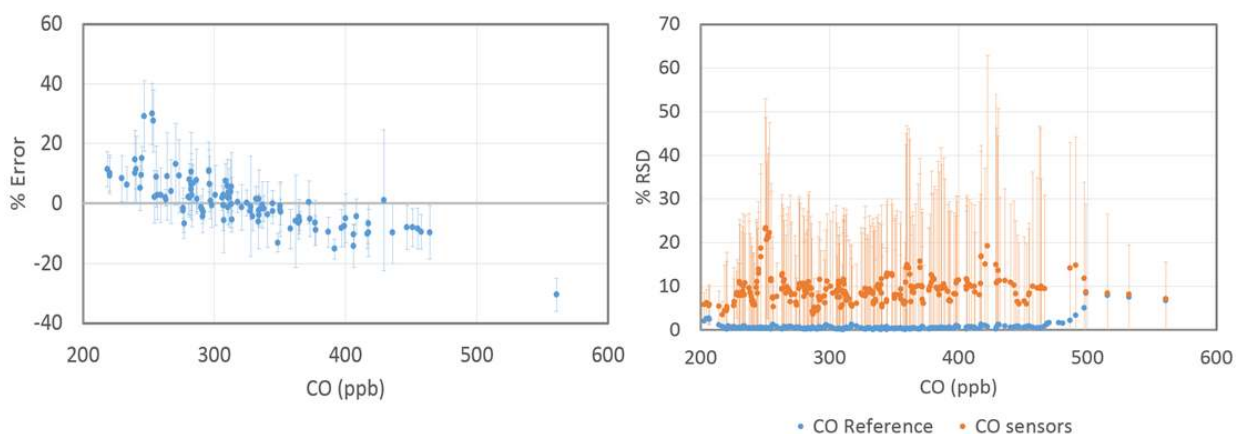


Figure 84. Evaluation of accuracy and precision of CO model with change in CO concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

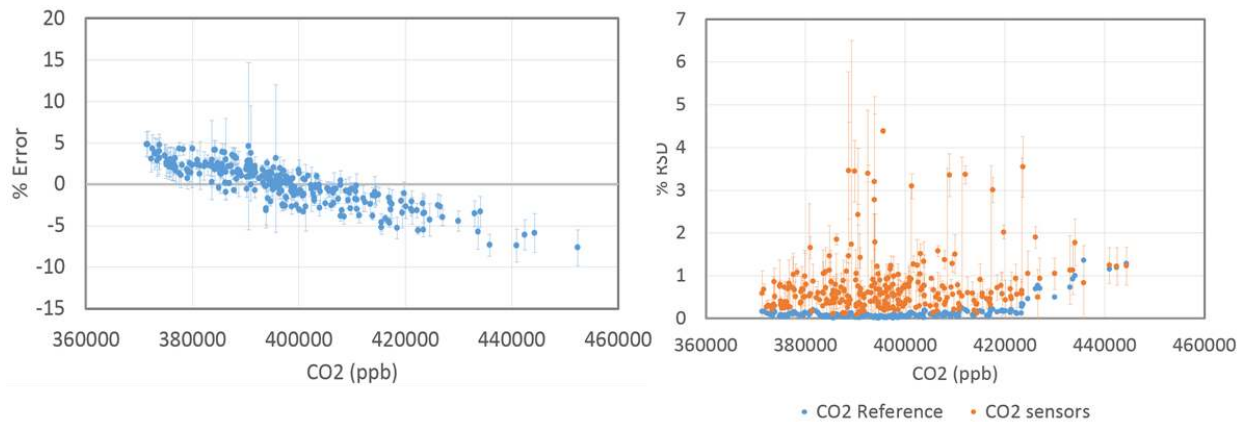


Figure 85. Evaluation of accuracy and precision of CO₂ model with change in CO₂ concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

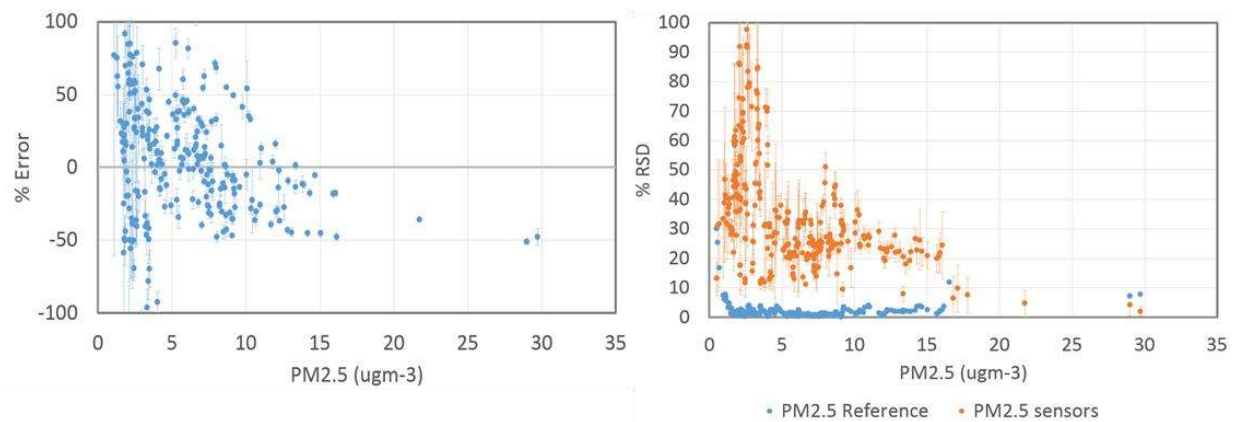


Figure 86. Evaluation of accuracy and precision of PM_{2.5} model with change in PM_{2.5} concentration. A) Accuracy is expressed as the percent difference from reference instruments as a function of pollutant concentration. Error bars represent variation between devices (SD between devices). B) Precision expressed as the relative standard deviation between replicate devices (n=4) as a function of pollutant concentration. Error bars represent variation between devices (SD between %RSD of devices). Original

data was collected between August 15, 2013 and May 10, 2014, sampled at 30-minute resolution, 6561 observations, 70% of data used. Data source: 2013-2014, AirSENCE Rev1.

8.5 Appendix 5: Wind impact analysis

It was found that at SOCAAR site, north-west wind was associated with higher traffic-related pollutant levels, while lower pollutant levels are generally associated with wind from the south (Figure 87). The air dynamics in the College street canyon is likely responsible for the increased level of air pollution associated with wind from north-west. Despite the fact that College street is directly south of the SOCAAR sampling site, wind from the south is deflected down the side of the building reaching the sampling port before interacting with the local traffic emissions. As a result, the influence of local traffic pollution is generally lower when the wind is blowing from the south. The consistent nature of this finding suggests that wind direction analysis is important for understand the impact of local sources on sensor response.

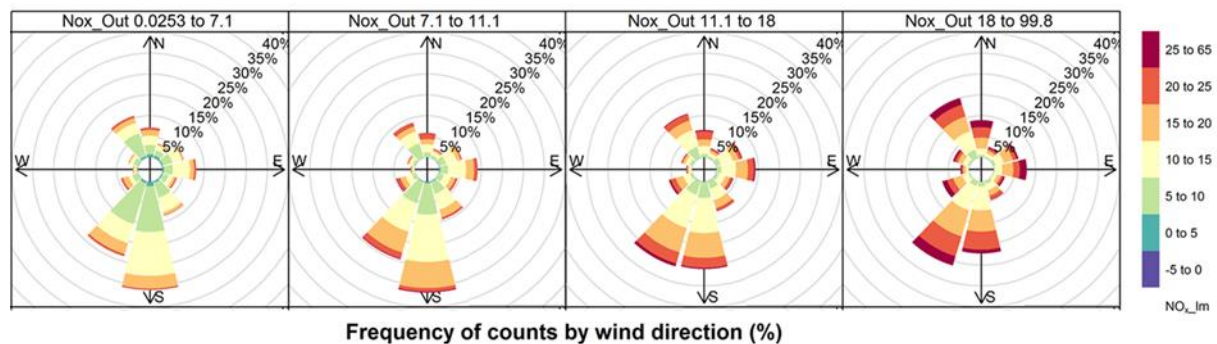


Figure 87. Assessment of the impact of wind direction on NOx concentration measured by AirSENCE devices and reference instruments. The graph is split into panels to show the dominant wind direction and speed for different ranges of reference NOx observations (Nox_Out): very low, low, medium and high. The NOx_lm colour scale indicates the NOx concentration measurements by the sensor device. Wind speeds are split into the intervals shown by the scale in each panel. The grey circles show the % frequency occurrence based on 1 min data. The plots show the proportion (here represented as a percentage) of time that the wind is from a certain angle and wind speed range. Mean values in each panel indicate the mean sensor NOx output for each panel (each bin of reference NOx observations). Original data was collected between September 20, 2015 and January 15, 2016, sampled at 1-minute

resolution, 139742 observations, 90% of data used. Data source: Multi-site validation study, Summer 2015, AirSENCE Rev2.

8.6 Appendix 6: CO/NO_x ratio analysis

In order to assess the similarity in source composition of SOCAAR site and other SCULPT sites, CO/NO_x ratios were then calculated for a set of representative peaks, which were coincident for the two pollutants. Background-subtracted data was used for calculating the ratios. Peaks were identified by calculating a smoothed first derivative and identifying when it crosses zero. In order to ensure that the selected peaks correspond to actual vehicle plumes, amplitude and slope thresholds were also applied to filter out the peaks that were too small or too wide. The peaks which were coincident for CO and NO_x were then selected and CO/NO_x ratios were found for each peak. CO/NO_x ratios were calculated for each site and ratio distributions were compared for different sites to test whether the contributions of NO_x and CO differ between sites and whether this difference is significant.

Before the CO/NO_x ratio was calculated for all sites, the average CO/NO_x ratios were compared for AirSENCE devices and reference instruments at the SOCAAR in order to establish the accuracy of sensor estimates. AirSENCE devices were found to underestimate the CO/NO_x ratios at by 30% on weekdays and 50% on weekends. The underestimation of the ratio was mostly caused by underestimation of CO concentration and was consistent over time. Therefore, the comparison of distributions of the ratios could still be useful for relative site comparison with respect to traffic pollution.

The differences in ratio distributions at central site and all other sites were found to be statistically significant (Kolmogorov-Smirnov p value < 0.05) but were especially pronounced for sites 12 and 18 (Figure 88). Based on this analysis, sites 12 and 18 were found to have highest CO/NO_x ratios. This suggests that these sites might be more influenced by gasoline emissions compared to diesel emissions. Site 10 was excluded from the analysis due to unreliable performance of CO calibration model.

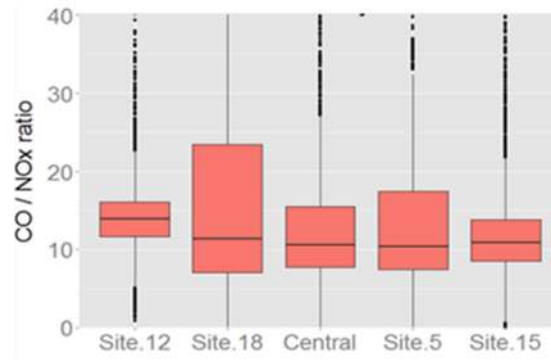


Figure 88. The comparison of CO/NO_x ratios (medians and distributions) for different sites (1 min average data). Original data was collected between August 8, 2013 and August 25, 2013, sampled at 30-minute resolution, 818 observations, 90% of data used. Data source: SCULPT Campaign, Summer 2013, AirSENCE Rev1.

