UNIVERSITY OF LEICESTER

## Theory and Practice of the use of Metal Oxide Semiconductor Pollution Sensors

by

Philip Peterson

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the College of Science and Engineering Department of Physics and Astronomy

February 11, 2020

### **Declaration of Authorship**

I, Philip Peterson, declare that this thesis titled, 'Theory and Practice of the use of Metal Oxide Semiconductor Pollution Sensors' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Pollution, all around.
Sometimes up, sometimes down.
But always... around.
Pollution, are you coming to my town?
Or am I coming to yours? Ha!
Were on different buses, pollution, but were both using petrol ... bombs!!!

- Rik, the People's Poet

### Abstract

The purpose of this work is to improve measurement capabilities for urban pollution monitors, making use of small, rapidly deployable and low-cost sensing solutions. A complete pollution monitoring instrument was designed and introduced in the course of this project, capable of supporting a wide variety of sensors. In this work, Metal Oxide Semiconductor gas sensors were used to measure concentrations of NO<sub>2</sub> and O<sub>3</sub> at an urban background site on University campus. These two gases are closely linked to traffic pollution and harmful to human health. The sensors were extensively tested in multiple experiments:

- Over one month the instrument achieved a residual standard error compared to the AURN reference (a chemiluminescence/UV fluorescence monitor) of  $23.6 \pm 3.9 \,\mu g \,m^{-3}$  for NO<sub>2</sub> and  $23.3 \pm 3.9 \,\mu g \,m^{-3}$  for O<sub>3</sub>.
- Analysis techniques were introduced to allow comparison of multiple calibrations over a long period. These could then be used to discard poor-quality calibrations, improving performance to a residual standard error of 22.9  $\pm$  5.3 µg m<sup>-3</sup> for NO<sub>2</sub> and 20.0  $\pm$  5.1 µg m<sup>-3</sup> for O<sub>3</sub>.
- Analysis techniques using separate calibrations for day and night were attempted, but produced very marginal results. Using multiple calibrations produced an improvement in residual standard error for NO<sub>2</sub> of  $3.4 \pm 6.2 \,\mu g \,m^{-3}$  and  $1.72 \pm 1.85 \,\mu g \,m^{-3}$  for O<sub>3</sub>. The effects of wind on sensor calibration were demonstrated, but an improvement in sensor performance through analysis could not be practically achieved.
- Using these analysis techniques, the effects of instrument design, such as active airflow, sensor "warm-up" time after power-up, and manufacturing variation in the sensors were investigated and two experiments using the unique characteristics of the instrument platform to examine the variation of pollutant gas on a small scale are described.

Some of these data analysis procedures are not specific to instruments making use of Metal Oxide Semiconductor gas sensors and may prove generally useful for other types of instrument.

### Acknowledgements

Name - Thanks

Roland - Quality bossing, trust.

Josh - Heroic support, advice and supervision.

James - Trust.

Tim, Mike - Getting me the job here in the first place, advice and help.

Fran, Eddie - Shelter, tolerance and wise friendship.

Kirsty - Crucial work on the sensor response; comradeship, laughter, schemes.

Amrita - Vital scientific and experimental work, particularly pertaining to wind analysis.

Martin - Adventurer, excellent instrument casing designer.

Antoine - Being a calming and patient colleague, CFD jiggerypokery.

Robin - A True Codewarrior.

Sams - Built data infrastructure that made the instrument work a million times better.

Chris - Coding drill sergeant, friend.

Rob, Alex, Dhan-Sham, Sefkan, Jasmine - Brave scientists who aided the development of this instrument by trying to use it.

Erika, Guillermo - Enriched this project with their skill, patience and experience. Harjinder - Consistently excellent advice, cakes.

Jasdeep - Doctorate Non Carborundum.

Ginga, Patrick - Carousing, guidance, sanity and affection.

Duncan - Tea, wisdom, tunes.

Natalia, Heather, Daniel, Marcus, Dmitri, Agnetta - Saw me at my worst, kept believing.

Nina - All the best.

Endless non-academic thanks to parents, friends, colleagues, owner of off-license.

## Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	iv
List of Figures	x
List of Tables	XXV
Abbreviations	xxix

1	Intr	oducti	on							1
	1.1	Overv	e <b>w</b>		 					1
	1.2	The H	uman Atmosphere		 					1
		1.2.1	Atmospheric struct	ure	 					2
	1.3	Airbor	ne pollution		 					5
		1.3.1	Pollutants in the u	rban environment	 			•		6
			1.3.1.1 Particulat	e matter	 					7
			1.3.1.2 Sulphur c	ompounds	 					11
			1.3.1.3 Volatile of	rganic compounds	 					13
			1.3.1.4 Carbon d	oxide	 					14
			1.3.1.5 Heavy me	$tals \ldots \ldots$	 					15
			1.3.1.6 Nitrogen	oxides and ozone	 			•		15
		1.3.2	Pollutant focus in t	his work	 			•		17
		1.3.3	Classifying environ	ments	 					18
			1.3.3.1 Indoor .		 			•		18
			1.3.3.2 Rural Am	bient $\ldots$	 			•		19
			1.3.3.3 Urban .		 					20
		1.3.4	Dynamics of the bo	oundary layer	 					21

		1.3.5	The interactions of air in cities
	1.4	Atmos	spheric chemical reactions of nitrogen oxides
		1.4.1	Nitrogen oxides and ozone
	1.5	Measu	ring nitrogen oxides
		1.5.1	Optical spectrometry
		1.5.2	Ground-based DOAS
		1.5.3	Airborne and satellite DOAS
		1.5.4	Chemiluminescence gas sensors
		1.5.5	Cavity Attenuated Phase Shift Sensors
		1.5.6	Advanced spectrometry techniques
		1.5.7	Metal oxide semiconductor sensors
		1.5.8	Electrochemical Sensors
		1.5.9	Diffusion tubes
	1.6	The re	ole of small sensors
	1.7	Resea	rch Objectives
	_		
2	Rev	view of	f low-cost gas sensors and instruments 50
	2.1	Metrie	cs of performance
		2.1.1	Accuracy requirements
		2.1.2	Uncertainty analysis
	2.2	Introd	luction to small low cost sensors
	2.3	Revie	w of available sensors $\ldots \ldots 58$
		2.3.1	Aeroqual SM50 O3
		2.3.2	SGX MICS series O3
		2.3.3	United Sens-IT
		2.3.4	Alphasense OX-series    62
		2.3.5	Citytech Sensoric O3-3E1F
		2.3.6	SGX MICS series NO2
		2.3.7	Alphasense NO2 series    65
		2.3.8	Citytech NO2- $3E50$
	2.4	Summ	hary: What sensors to choose?
	2.5	How I	MOS sensors work
	2.6	Revie	w of instrument designs 74
		2.6.1	$AQMesh \dots \dots$
		2.6.2	Cairpol CairClip
		2.6.3	$Aeroqual Series 500 \dots 78$
		2.6.4	United SENS-IT and ETL-3000
		2.6.5	mPod
		2.6.6	ELM
		2.6.7	$#AirQualityEgg \dots $
	2.7	Summ	nary

3	Nov	el calibration and data analysis techniques for a bespoke en-	
	vire	amental sensor 8	5
	3.1	Introduction to SOGS and Zephyr	7
		3.1.1 Instrument architecture	9
		3.1.2 Other instrument hardware $\ldots \ldots \ldots \ldots \ldots \ldots $ 9	1
		3.1.3 Amplifier development	2
		3.1.4 Timeline of deployments	4
	3.2	Review of calibration techniques	5
		3.2.1 Background of linear calibration equations 9	7
		3.2.2 Cell Calibration	9
		3.2.3 Calibration strategy: controlled cells or co-location? 10	1
	3.3	Outdoor calibration	5
		3.3.1 Physical configuration of the AURN site	7
		3.3.2 The long-term experiment	8
		3.3.3 Raw sensor data $\ldots \ldots 10$	9
	3.4	Preliminary analysis	6
		3.4.1 Candidate equations $\ldots \ldots 11$	7
		3.4.2 Optimal calibration time $\ldots \ldots 12$	0
		3.4.3 Combined long-term data analysis	4
		3.4.4 Issues with long term experiment data $\ldots \ldots \ldots \ldots \ldots 12$	7
	3.5	Comparison of fitting equations	0
		3.5.1 Discussion $\ldots \ldots 13$	1
	3.6	Calibrated sensor performance	2
		3.6.1 Discussion $\ldots \ldots 13$	4
	3.7	Selecting good fits $\ldots \ldots 13$	7
		3.7.1 Discussion $\ldots \ldots 14$	1
	3.8	Summary	3
4	Une	erstanding pollution sensors in changing environmental con-	
	diti	ns 140	6
	4.1	Overview of advanced fitting techniques	7
	4.2	Correcting for day/night chemistry	0
		$4.2.1  \text{Method}  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  $	0
		4.2.2 Results and discussion	2
		4.2.3 Summary of outcomes	7
	4.3	Transitional fits	7
		$4.3.1  \text{Method}  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  $	9
		4.3.2 Results	0
		4.3.3 Discussion $\ldots \ldots 16$	1
		4.3.4 Summary of outcomes	4
	4.4	Wind direction and calibration	6

		4.4.1 Experimental conditions	166
		4.4.2 Analysis	167
		4.4.3 Results	168
		4.4.4 Further investigation	170
		4.4.5 Summary	174
	4.5	Timeshift error analysis	174
		4.5.1 Method	174
		4.5.2 Discussion	176
	4.6	Summary	176
5	Cha	aracteristics of MOS sensors and their implications for instru	-
	mei	nt design	178
	5.1	Active airflow - yes or no?	179
		5.1.1 Results $\ldots$	180
		5.1.2 Discussion $\ldots$	181
		5.1.3 Summary of findings	183
	5.2	Warm-up experiments	183
		5.2.1 Experimental method	186
		5.2.2 Results and discussion	187
		5.2.3 Summary	188
	5.3	Manufacturing variation in MOS sensors	190
		5.3.1 Method $\ldots$	190
		5.3.2 Discussion $\ldots$	192
	5.4	Summary	193
6	Exp	perimental study on spatial variation of pollutants	195
	6.1	The arm's length scale	196
		6.1.1 Experimental setup	197
		6.1.2 Voltage correlation analysis	198
		6.1.3 Fitting analysis	200
		6.1.4 Residuals and atmospheric variation	202
		6.1.5 Summary	207
	6.2	The garden scale	209
		6.2.1 Method	209
		6.2.2 Results	211
		6.2.3 Discussion	218
	6.3	Summary	223
7	Cor	nclusions and further work	224

 $\mathbf{232}$ 

	A.1	Introduction	232
	A.2	Revision history	234
		A.2.1 Version 0.1	234
		A.2.2 Version 0.2.3	235
		A.2.3 Version 0.2.5	237
		A.2.4 Version 0.2.11	239
	A.3	Overview of most recent design	241
		A.3.1 Major component description	241
		A.3.2 Supporting component description	244
В	Var	iation in temperature and humidity sensors 2	246
	B.1	Experimental setup	246
	B.2	Results	247
С	Cas	e study: Random forest machine learning 2	251
	C.1	Demonstration	253

#### Bibliography

256

# List of Figures

1.1	Diagram of the changes in Earth's atmospheric temperature (bot-	
	tom axis, solid line) and pressure/density (top axis, dashed lines.)	
	with height. Four atmospheric layers are labelled here, alongside	
	the temperature inversions that define the different layers. The Ex-	
	osphere starts and extends well beyond the vertical scale of this dia-	
	gram. Pressure and density both decrease with height, but not quite	
	in the exponential way one would expect mathematically, based on	
	the hydrostatic equation. The different layers are picked out with	
	an orange tint. Adapted from [2, fig 2.6]	2
1.2	A gravimetric sensor. Air enters through A, a size-selective inlet	
	which rejects particles above a certain mass, and is drawn through	
	the filter $\mathbf{B}$ by the pump $\mathbf{C}$ at a constant rate	8
1.3	Schematic of a Beta Attenuation Monitor. Air enters through the	
	size selective inlet $\mathbf{A}$ , and is pulled through a porous ribbon by the	
	pump <b>B</b> . The ribbon drawn from reel <b>C</b> onto <b>F</b> , across the airstream	
	at a steady speed, and after exposure it passes under a beam of beta	
	radiation from a radioactive source $\mathbf{D}$ . The more particle mass is	
	stuck in the ribbon, the more attenuated the radiation will be when	
	it reaches the detector $\mathbf{E}$	9
1.4	Schematic of a Tapered Element Oscillating Microbalance. Air en-	
	ters through the size selective inlet $\mathbf{A}$ , before splitting in two at	
	<b>B</b> , to create a bypass that is used to regulate the amount of air	
	passing through the sensor element, a tapered glass cylinder C.	
	One end of this cylinder is fixed to the instrument, but the other	
	is free to vibrate at its natural frequency. This frequency is contin-	
	ually measured electronically <b>D</b> , and as particles accrete onto the	
	cylinder, their mass affects this frequency, like a singing wineglass	
	changing its note as your pour in more liquid. The valves <b>E</b> control	
	the amount of air flowing through the sensor and bypass line to the	10
	exhaust pump $\mathbf{r}$	10

Schematic of a Shinyei PPD42 nephalometer. Air beneath the hous- ing <b>A</b> is heated by the resistive element <b>B</b> , drawing air up through the sensor. An infrared LED <b>C</b> shines into this air, and the dif- fuse reflected glow from particles within the rising stream <b>D</b> will be	
detected by the photodiode $\mathbf{E}$ , which is obscured from the LED Schematic of a UV fluorescence sensor. Air $\mathbf{A}$ is pumped through the sensor and exposed to light pulses from lamp $\mathbf{B}$ , at a frequency capable of exciting SO <sub>2</sub> molecules. The photons these re-emit pass	11
Schematic of a 2D Gas Chromatograph. Air is pumped into the sen- sor and split into two streams at <b>A</b> . The airstreams pass through valves which controls the pressure in the 'columns' <b>C</b> and <b>D</b> . The two columns contain different granular material, through which dif- ferent gas species will percolate at different speeds. At the end of the columns is a Flame Ionization Detector (FID) <b>E</b> , which ignites the gas output (the 'eluate') in a hydrogen flame and detects the in- tensity of the ionized combustion products using electrostatic plates	12
<b>F</b>	13
erly calibrated photodiode <b>C</b>	15
haust plume in those environments. Adapted from [83] Isopleths giving the net rate of ozone production (in part-per-billion per hour) as a function of VOC (part-per-billion carbon) and NOx (part-per-billion) concentrations, for mean summer daytime meteorology and clear skies. The curved lines represent production rates of 1 2.5 5 10 15 20 and 30 ppb/h. Image credit[86]	23
Diagram of a generic grating spectrometer. Collimated light enters aperture <b>A</b> and is focused by mirror <b>B</b> onto the diffraction grating <b>C</b> . The diffraction pattern is refocused by <b>D</b> onto an array of pho- todiodes (or a CCD) indicated by <b>E</b> . Signals from the detector <b>F</b> must be interpreted by computer to determine from the diffraction	20
pattern which frequencies are present	33 33
	Schematic of a Shinyei PPD42 nephalometer. Air beneath the housing <b>A</b> is heated by the resistive element <b>B</b> , drawing air up through the sensor. An infrared LED <b>C</b> shines into this air, and the diffuse reflected glow from particles within the rising stream <b>D</b> will be detected by the photodiode <b>E</b> , which is obscured from the LED. Schematic of a UV fluorescence sensor. Air <b>A</b> is pumped through the sensor and exposed to light pulses from lamp <b>B</b> , at a frequency capable of exciting SO <sub>2</sub> molecules. The photons these re-emit pass through the optical filter <b>C</b> and are detected by the photodiode <b>D</b> . Schematic of a 2D Gas Chromatograph. Air is pumped into the sensor and split into two streams at <b>A</b> . The airstreams pass through valves which controls the pressure in the 'columns' <b>C</b> and <b>D</b> . The two columns contain different granular material, through which different gas species will percolate at different speeds. At the end of the columns is a Flame Ionization Detector (FID) <b>E</b> , which ignites the gas output (the 'eluate') in a hydrogen flame and detects the intensity of the ionized combustion products using electrostatic plates <b>F</b>

1.13	Diagram of a DOAS instrument taking a measurement. Light from	
	the sun A scatters as it passes through atmospheric gases B, and	
	the light that is scattered along the line of the instrument aperture	
	C enters the instrument D. The situation in reality is complicated	0.4
	by the curvature of the atmosphere, refraction and multiple scattering.	34
1.14	(a) Cityscan atop a root in London. (b) Example of data from a	
	CityScan deployment in London[96], a heatmap of $NO_2$ concentra-	~ ~
	tion. The "hotter" colour corresponds to higher concentrations	35
1.15	Open path DOAS using a downward-pointing spectrometer. Light	
	from the sun <b>A</b> is reflected off the ground <b>B</b> and into the spectrom-	
	eter on the air or spacecraft C. The gas concentration E within the	
	column <b>D</b> formed by the path of the light from the ground to the	
	spectrometer can be retrieved	36
1.16	Airborne air quality mapper, the precursor to ANDI, aboard an aircraft operated by BlueSky, with data from over Leicester[96].	37
1.17	Diagram of the operation of a chemiluminescence sensor. Intake air	
	is split at <b>A</b> , one stream of which is put into a catalytic converter	
	<b>B</b> that reduces $NO_2$ to $NO$ . The values at <b>C</b> control which stream	
	enters the chamber ${\bf E}$ and reacts with ozone produced with the UV $-$	
	lamp at <b>D</b> . The photons emitted by this reaction are detected with	
	the light sensor $\mathbf{F}$ , and the gas is then exhausted through the pump	
	<b>G</b>	38
1.18	A typical modern commercial chemiluminescence monitor, from	
	Ecotech.	39
1.19	Diagram of a CAPS instrument. The air sample is drawn through	
	the particle and humidity filter A to the cavity B. A lamp C shines	
	through this cavity, bouncing between the mirrors <b>D</b> before leaking	
	through to the detector <b>E</b> . The cavity is drained using the pump $\mathbf{F}$ .	
	Representation of the source signal (intensity vs. time) at C and	10
1 00	output signal at $\mathbf{F}$ are also shown	40
1.20	Diagram of an FTIR gas spectrometer. The beam from light source	
	A is split by C, and one of the resulting beams bounces from the	
	fixed mirror <b>B</b> . The other shines through the sample chamber $\mathbf{D}$ ,	
	and bounces on the movable mirror <b>E</b> . The beams merge again	
	at the splitter and are detected by the sensor $\mathbf{F}$ . The computer	
	G controls the mirror position, and uses a Fourier transform to	
	reconstruct the spectrum of light from the position-dependent in-	11
1 01		41
1.21	Diagram of a DDULAS cavity arrangement. The light source $\mathbf{A}$	
	since since $\mathbf{D}$ and $\mathbf{D}$ , and is then repeatedly reflected through the sample shamper $\mathbf{C}$ . The mirrors are not perfectly reflected and	
	photons from the source light can eventually escape to the spec	
	tromotor <b>D</b>	/1
		41

1.22	A laboratory Michelson-equipped BBCEAS system built by Uni-	
	versity College Cork in Ireland's Laser spectroscopy group[122]	42
1.23	Diagram of a MOS sensor. The air outside the housing <b>A</b> filters	
	through the aperture to the semiconductor element $\mathbf{B}$ , which is	
	heated by $\mathbf{C}$ .	43
1.24	Diagram of an electrochemical gas sensor. The sample air $\mathbf{A}$ passes	
	through the grille to the working chemical electrode <b>B</b> . Isolated	
	from this air is a reference electrode C which is used to compensate	
	for temperature changes and leakage currents. The electrodes feed	
	that is a waltage is applied agrees them through $\mathbf{F}$	11
1.95	- that is, a voltage is applied across them - through E.	44
1.20	A NOw diffusion tube cold by Credite International	45
1.20	A NOX diffusion tube sold by Gradko International	40
2.1	Data from the University of Leicester AURN site showing concen-	
	trations of $NO_2$ and $O_3$ through 2016, per hour of the day. The	
	central line is the median value, the shaded area represents the first	
	and third quartile, and the dashed lines represent the 5th and 95th	
	percentile. Maximum 95th percentile $NO_2$ and $O_3$ for this site in	
	2018 was 52.0 $\mu$ g m <sup>-3</sup> and 83.27 $\mu$ g m <sup>-3</sup> respectively[148]	54
2.2	SM50 O3 daughterboard. The grey cylinder on the left is the sensor	-
	element.	58
2.3	MICS OZ-47 demonstration board	60
2.4	MICS legged sensor	60
2.5	MICS surface mount sensor	61
2.6	Sens-IT module, with inlet and outlet on the top and bottom	62
2.7	Alphasense A-series	62
2.8	Alphasense B-series	63
2.9	O3-3E1F sensor, mounted on a circuit board	64
2.10	MICS surface mount package	64

2.11	Diagram of the various types of interaction between atmospheric gases and a MOS sensor surface. In the leftmost region the sensor is unpowered (and exhibits the base resistance). The three other	
	regions of the diagram describe different processes that actually oc- cur simultaneously to varying degrees. The sensor's output is the	
	resistance across the whole of the sensor material, which forms a resistor network with contributions from both the bulk and surface regions (although the non-sensitive surface will have similar prop-	
	the wide variation in base resistance between individual sensors of the same type, as the random nature of the surface geometry means	
	an equally random network of resistances. This diagram is a two dimensional representation of a three dimensional material; in an actual sensor, the sensitive region is spread into the surface with a	
	published by the author in Peterson et al. [202]	73
2.12	The interior structure of the CairClip instrument with an electro- chemical sensor. Image credit [206].	76
2.13	Aeroqual Series 500 handheld pollution sensor. Image credit Aero- qual [166].	78
2.14 2.15	Weatherproof enclosure of the United ETL-3000 instrument, which makes use of the SENS-IT sensor. Image credit S.R.L. [179] The interior structure of the mPod_Image credit Piedrahita et al.	79
2.10	[50]	80
2.16	Air quality egg exterior. Image credit Chantrell [216]	81
3.1	Map of Leicester city, with north is directly upward. Districts are labelled in orange and divided by white borders. The two AURN stations are labelled in blue, with the University of Leicester station	
3.2	In the Victoria Park district	86
	through 2016, per month. The central line is the median value, the shaded area represents the first and third quartile, and the dashed	0.0
3.3	<ul><li>(a) A schematic of the main board of SOGS with functional regions labelled.</li><li>(b) SOGS daughterboard example configuration.</li><li>This diagram shows base board, an expansion board (through the</li></ul>	88
	digital Raspberry Pi header) and instrument boards (through the analog headers). Multiple boards can be stacked through the digital	00
	header, and up to four different analog instruments can be connected.	89

3.4	Image of the layout of sensors on a SOGS MOS instrument board. Visible are the (a) set of sensors; the second set of sensors (b) is on the reverse side. The MICS4514 contains two different sensor	00
3.5	Cutaway of the prototype Zephyr enclosure, showing layout of vital components and interfaces	. 90
3.6	Diagrams of SOGS amplifier circuits	92
3.7	(a) MICS-4514 Oxidising gas sensor element nonlinear response to nitrogen dioxide.[190] (b) MICS-4514 Reducing gas sensor element response to various gases. $R_s$ is the sensor's actual resistance, com-	
3.8	pared to the base resistance $R_0.[190]$	. 98
3.9	to the reference for the high concentrations. Typical levels of, $NO_2$ in urban environments are not often higher than 100 µg m <sup>-3</sup> [149], although peaks can be ten times as high. For concentrations less than this, both normal and inverse voltage can be seen as approxi- mately linear. Graph (b) shows the data for inverse voltage at low concentrations	. 100
	between $10-100 \mu g \mathrm{m}^{-3}$ , so the interference from $CO_2$ is significant for this sensor type[222].	. 103
3.10	Isometric diagram of the relative positioning of Zephyr sensors near the roof of the AURN station. The lip of the roof is roughly 3	
3.11	meters from the ground	. 105
	labelled in blue. North is directly up in the image	. 107

3.12	Data from the LTE dataset after zero voltage readings have been re- moved, showing typical voltage outputs from the OXa and O3a sen- sors alongside temperature (Temp.) and relative humidity (%RH)
	information, and the $NO_2$ and $O_3$ readings of the AURN reference
0.10	instruments
3.13	Scatter plots of the resistance response of the oxidising (OX) and ozone (O3) sensors to target gases, for data taken within 45 and 55% relative humidity and 23 and 27°C. Note how, in the left panel, the NO <sub>2</sub> data for sensors c and d occur in two distinct domains, offset
	vertically by one to one and a half ohm
3.14	Correlation meander over time for long term experiment sensors, both O3 and OX. Two gaps that indicate obvious changes in cor- relation are marked with dashed red lines, and have been labelled with possible explanations given in the text. The y units are di- mensionless, given that they are the ratio between one variable and the average of several of the same kind. The meander's value is
	thousandth 112
3.15	(a) Scatter plots showing relation between a, b, c and d sensor
	resistances for both OX (upper right graphs) and O3 (lower left). The scale is the same on every plot of the same type of sensor: 1783 to 1786 ohms for the OX sensors and 1781 to 1786 for the O3 sensors. The scale is less important than the shape of the different coloured patches, and the colours correspond to different time periods, shown in (b). These time period colours are delineated by the end of the
	brownouts in May 2016 and then the power cuts in November 2016. 114
3.16	Scatter showing resistance of sensor O3b vs O3d during the long term experiment, colour coded in the same way as 3.15(b),
3.17	A month-long timeseries from the long term experiment between 17th of March and 17th of April 2016, with predictions from the MOS sensor data using equations 3.16 and 3.20 for NO <sub>2</sub> and O <sub>3</sub> respectively. The black lines in the NO <sub>2</sub> and O <sub>3</sub> sections are ref- erence data from the AURN station instruments. The week-long calibration period is shown as a red box on the humidity timeseries 120
3.18	Various metrics of quality of calibration vs length of calibration
5.10	period, for all four sets of sensors. $NO_2$ and $O_3$ are fitted using the "inverse NO2 linear O3" equation 3.17 and the "inverse" equation 3.20 respectively. The different coloured lines correspond to OX and O3 sensors a, b, c, and d, and the best point on each quality curve is marked with a circle. The metrics of quality compare the predictions of the sensors with those of the AURN station during
	the validation period, a week directly after the calibration 121

3.19	Visual representation of a calibration taken over a weeklong part of	
	a larger data series (dark orange) being used to create a weeklong	
	prediction (light orange), from which quality metrics can be taken	
	by comparing the prediction with the reference instrument. The	
	grey line represents the underlying data being produced by both	
	the sensor being calibrated and the reference.	. 124
3.20	Visual representation of set of predictions over the whole reference	
	timeseries. $t=0$ at the beginning of the particular calibration section	
	that produced the predictions	. 125
3.21	The grey hexagons represent ratings, the darker the hexagon the	
	better the rating. Effectively the closer the rating is to the calibra-	
	tion period, the better it is likely to be	. 125
3.22	Collection of ratings arrays for different calibration sections. This	
	concept is further illustrated in figure 3.25.	. 126
3.23	Shifting the ratings arrays to align their timestamps.	. 126
3.24	Averaging the ratings together. The red curve is an example of a	
	series of averaged ratings.	. 126
3.25	Explanation of the time axis of CAFPOLD plots. The "source"	
	axis shows arbitrary predicted and reference gas concentration data.	
	The red box shows the section of the source data used for fitting a	
	calibration equation. The "CAFPOLD" axis is the graph of a given	
	quality metric over time. The purple rectangles on the source axis	
	show the time periods over which the quality metric is calculated.	
	The value of that metric corresponds to a point on the CAFPOLD	
	axis indicated by a circle. Point <b>A</b> is on negative time because it	
	represents a prediction made from calibration data that occurred in	
	the future. The blue lines at $t_c$ on the CAFPOLD graph bracket	
	the region of points where the quality metrics are calculated for	
	predictions partially during calibration. For example the first half	
	of the rating for point <b>B</b> covers data just before calibration, the	
	latter half covers some data from within the calibration period.	
	Point <b>C</b> shows the point on the CAFPOLD graph corresponding	
	to quality metrics calculated directly after the calibration period	
	ended. It should be noted that for this graph a single calibration	
	on a single section of reference data is displayed. A real CAFPOLD	
	analysis takes the average of many such calibrations.	. 127

- 3.27 CAFPOLD analysis comparing different fitting equations. The vertical blue line signifies a week from calibration, which is at zero on the X axis. There are three plots for each gas for the three quality metrics. Gaps in the plot traces correspond to missing data or sections where the fit could not produce a prediction. Once again the specification thresholds are denoted with black horizontal lines,  $20 \,\mu g \,m^{-3}$  for the solid line,  $40 \,\mu g \,m^{-3}$  for the dashed line. . . . . . . 131

- 4.1 Example of a CAFPOLD analysis using two weeks of calibration data with the diurnal sectioning. The output for each sensor is shown as a separate trace. Note that both axes of the graphs are different to that of previous CAFPOLD analysis diagrams, spanning one and a half months of time in the x axis and different scales for the *RSE* and *FE* in graphs a) b) and d) e). Sensors c and d exhibit poor performance, particularly around +1 month in the *RSE* graph a). The blue vertical lines are the period of calibration. . . . . . . . 154

4.2	Graph of CAFPOLD results for $NO_2$ and $O_3$ . The graph uses the	
	averaged predictions of the sensors used in the experiment. Each	
	colour of line corresponds to a different sectioning method. Once	
	again the vertical scales here are different to the other CAFPOLD	
	analyses in this work. The horizontal black line indicates $20 \mu g  m^{-3}$	
	target performance.	. 155
4.3	CAFPOLD graph of difference between sectioning methods and the	
	baseline 14-day calibration. Each colour of line corresponds to the	
	predictions using a particular sectioning method minus the baseline	
	prediction. The scale is zoomed in to make the difference in methods	
	clearer	. 156
4.4	Timing diagram for the transitional fit plots in this chapter. The	
	time represents the beginning of the block of starting calibration.	
	evaluation, and ending calibration. Each point on the output qual-	
	ity plot shows the quality achieved during the evaluation period for	
	the respective time	. 161
4.5	Evolution with time of residual standard error, fractional error and	
	Pearson correlation coefficient for the three fitting methodologies.	
	for prediction $NO_2$ concentration. The two lines in each colour are	
	for sensor a and b; sensors c and d performed very poorly and are	
	not shown here. The dates are all in 2016.	. 162
4.6	Evolution with time of residual standard error, fractional error and	
	Pearson correlation coefficient for the three fitting methodologies,	
	for prediction $O_3$ concentration, with multiple traces for sensors a,	
	b, c and d	. 163
4.7	Wind vector data, using the two averaging methods. Wind direction	
	is given in terms of source, so a northerly wind blows air from north	
	to south. The positive Y axis of the graph points north	. 168
4.8	Scatter plot of wind vectors, coloured according to the difference	
	between predicted and actual concentrations of a) $NO_2$ and b) $O_3$ .	
	Sensor a from both cases is the one shown. Anomaly refers to the	
	predicted gas concentration minus the reference concentration. The	
	segments drawn on the diagram each contain the same number of	
	points, and indicate density of wind vector points that might not	
	be apparent from the points alone.	. 169

4.9	Violin plots of difference between predicted and actual data, divided according to the prevailing wind direction. For each violin plot, the thicker the band between the vertical axis and the edge of the violin, the more data points lie within the region quantified by the vertical axis. There are multiple violins, each one representing a 20 degree segment of the wind rose. Some violin strips are very narrow and represent very few overall data points in this wind direction, the lowest having five. The (green) left and (purple) right sides of each "violin" represent the two different sensors, a and b. The darker coloured beads indicate the mean of the anomaly for that sensor for wind direction within the remostive angle	171
4.10	which direction within the respective angle	
4.10	As above, but for $O_3$ . A wind scatter plot for $NO_2$ sensor a, overlain with boundaries chosen such that the data is divided into six segments, excluding the innermost quintile of wind speed data, and each containing the same number of points. The <i>RSE</i> value of each segment is superimposed. The colour of the segments emphasises this (ranging through green	. 1/1
4.12	for the lowest <i>RSE</i> to red for the highest one)	172
	inverse $NO_2$ equation, no fit can be made	173
4.13	Results of the timeshift analysis. The x data is in hours, with zero	
	representing zero timeshift	175
5.1	Experimental setup for investigating aspirated (on the left) vs. pas- sive sampling (right side), on the AURN station	180
5.2	Timeseries and scatter of active airflow instrument (red) and passive airflow instrument (blue). The timeseries plots these predictions alongside the AURN reference (black). The graph on the right is a scatter plot of the same data in the same colours, with the 1:1 line marked as a dashed black line.	181
5.3	Graphs of change in performance metrics for the active and passive instruments, when the active airflow instrument has a moving av- erage applied to it. The period of the moving average window is shown in the horizontal axis, and ranges from zero (ie. no averaging window) to thirty-six hours. The horizontal dashed black line is the	
	passive instrument's performance statistics	182
5.4	Example curve of net gas adsorption after a power supply interrup-	
	tion event	186
5.5	A timeline of the power supply interruption experiment, which took place over multiple days in December 2016. The labels for the peri- ods during which the power supply was interrupted are referenced	
	in the results section	187

5.6	Fitted curves after power supply interruption for four different sen- sors, two each of oxidising (OXa and OXb) and ozone (O3a and O3b), in a stable atmosphere. The vertical axis of each graph is the increase in voltage either side of the disconnection period. The baseline that the sensor output settled to when left powered for at least an hour was set at zero volts (see figure 5.5. The blue curves are the raw voltage curves, some of which show the hysteresis of the sensor toward the beginning each curve (particularly for OXa and OXb). The red lines are continuations of the voltage curves to the time at which the power supply was reconnected, ignoring the hysteresis. These terminate in red circles at the projected voltage difference at the moment the sensor was reconnected. The dashed green line is an exponential curve fitted to these points showing the likely voltage curve when the power is disconnected and the sensor is cooling. The solid green line is the unpowered baseline voltage	
57	inferred from this fit. The grey lines either side of the likely voltage curve delineate the 95% confidence interval of the exponential fit	. 189
ə. (	a) Violin plot of PCC between a and b sensors. The mean correla- tion has been marked with a white dot. b) Gradient of least-squares fit between a and b sensor voltages, with a value of one implying a 1:1 correlation. Violin plots are broader where more results are in the region indicated by the vertical axis.	. 191
6.1	Layout of the multi-sensor adjacency experiment. Blue lines indi- cate dimensions. Positions are given relative to the leftmost sensor's intake. The instruments (apart from the LTE, long term experi- ment) are demarcated by a name of the form DCAxx, where xx are two numerals. The red crosses indicate sensors that were found to not produce accurate results - see later in this section.	. 197
6.2	This scatter grid plot shows the correlation from all six functional sensors (that is, twenty four sensor elements, twelve each for OX and O3 MOS sensors). The bottom left portion of the graph shows the O3 sensors, the top right are OX. Pairs of a and b sensors are adjacent to each other and are shaded the same colour, helping emphasize the blocks of four scatter plots. The axes are normalised so that the center of each plot axis is the mean voltage of the sensor for that axis. The scale is individual to each scatter, and is such that both axes of the scatter cover a range of voltage equal to twice the sum of standard deviations for both of the sensors indicated by	
	the axes.	. 199

6.3	This scatter grid plot, similar to 6.2, shows the correlation between the predictions of the twenty four sensors - in the upper right for $NO_2$ , in the lower left for $O_3$ . As with figure 6.2, the scale of the individual scatter plots centers around the mean of the predicted data and extends to the sum of the standard deviations. This keeps the data centered but does not transform it or change the angle of the individual plots. Colour is used to visually group a and b sensors from the same instrument together. These scatter plots are between
6.4	Zephyr-prototype sensors, not the AURN reference instruments 203 Residual between sensors for $NO_2$ . Each point is the standard devi- ation of the residual standard error between the labelled sensor and all the others in the experiment. The x axis arranges the instru- ments in their approximate position on the AURN station, with results for sensors a and b for each instrument (in red and blue) shown next to the intra-instrument comparison between a and b(in grey). The separate axis on the left of the diagram shows the aver- age of the inter-instrument comparison. The error bars are to the
	first standard deviation (63% confidence). $\ldots \ldots \ldots \ldots \ldots 206$
6.5	As above, but for $O_3$
6.6	Roving instrument cart beside the Leicester University AURN sta-
~ -	tion. Only the instrument at 150 cm elevation is used in this analysis.210
6.7	Path and relative concentrations of the roving sensor experiment superimposed on an aerial image of the ground near the AURN station. The average distance moved between measurements is 6.93
6.8	$\pm 2.05 \mathrm{m}$
6.0	Timeseries of predicted concentrations for both NO <sub>2</sub> and $O_2$ 213
6.10	Scatter of predicted concentrations, for both $NO_2$ and $O_3$ 215 Scatter of predicted concentration vs. road proximity for $NO_2$ sensors a and b. The road proximity factor is a normalized, dimension- less figure obtained by dividing the road proximity equation (6.6) by its minimum value 216
6.11	Scatter of wind direction during roving sensor experiment, coloured according to a) NO <sub>2</sub> or b) O <sub>3</sub> concentration from sensor a within the 150 cm instrument. The colour indicates the gas concentration measured by the sensor for a particular wind direction. A random number (between -0.05 and +0.05 m/s) was added to the x and y direction of each wind measurement to make distinct the different
6.12	$\begin{array}{llllllllllllllllllllllllllllllllllll$
	of the road, for the given levels of influence from the road. $\ldots$ . 219

6.13	a) is a scatter showing of oxidant gases verses NOx taken at a six urban and suburban sites in and around London, averaged over the night time, hours, during two Neverthers in 1008 and 1000. Dis	
	ture credit to [89] The values are in part per billion the trans-	
	lation of which to micrograms per meter cubed depends on the	
	molecular weight of the gases in question. b) shows the scatter	
	graph of oxidant gases $(NO_2 \text{ plus } O_3)$ verses NOx taken at the	
	University of Leicester AURN station, hourly, during November	
	in 2015. The trendline and y intercept are marked with a solid	
	and dashed line respectively. The fitted trendline has the equation $[0-1] = 0.252[NO-1] + 65.54$	001
614	[OX] = 0.355[NOX] + 05.54.	. 221
0.14	the roving sensor experiment, for each of the measured gases, and	
	the total oxidant level, compared with the CFD simulated contri-	
	bution of gas coming from the road, a dimensionless concentration	
	factor.	. 222
7.1	Timeseries of reference (from FTIR instrument) and MOS predicted	
	gases, $CH_4$ and $N_2O$ respectively. The units are preserved from the	
	FTIR - part per million and part per billion	. 229
A.1	Version 0.2.3 layout. The Raspberry-Pi header is in the lower right.	
	The major subsystems are identified by coloured blocks.	. 235
A.2	Version 0.2.5 layout. The major systems are identified by coloured	
	blocks. The power subsystem is split into three locations on this	
	design.	. 237
A.3	Version 0.2.11 layout	. 239
A.4	Schematic diagram of SOGS 0.2.11c. The signals and power wires are represented by lines coloured to indicate their function. The	
	grey boxes are subsystems that may be comprised of multiple elec-	
	tronic components. Circles are terminals that can accept single	
	external connections or testing probes, rounded bars are headers	
	can be soldered over to change the circuit's functionality. A dashed	
	outline indicates an external component.	. 241
	r	

B.1	Grid of scatter plots between six sets of Zephyr temperature and	
	humidity sensors. The lower left grids show scatters between humid-	
	ity sensors in different instrument housings, the upper right grids	
	are for temperature sensors. The $PCC$ between pairs of sensors is	
	written at the bottom of the grid axes. Above each grid square is	
	the mean and $95\%$ confidence of the difference between the X and	
	Y sensor. The scale of the grids is normalised such that the aver-	
	age of the data set is in the center, and the upper and lower axis	
	boundaries are at points equal to this mean plus twice the standard	
	deviation of the data. The scale is identical among all temperature	
	and humidity grids respectively.	. 248
B.2	Plot of <i>PCC</i> between sensors verses distance between sensors being	
	compared, for both humidity and temperature. The diamond mark-	
	ers are for comparisons involving DCA82, and are omitted from the	
	calculation producing the trendline	. 249
$C_{1}$	Results of CAEPOLD analysis using random forest to predict NO <sub>2</sub>	
$\bigcirc$ .1	and $\Omega_{\rm c}$ concentration	253
		. 200

## List of Tables

1.1	Summary of different $NO_2$ measurement techniques. Unmarked concentrations are in units of $\mu g m^{-3}$ . All the in-situ instruments except for diffusion tubes are capable of sacrificing time resolution to improve accuracy and range.	46
2.1	Table summarising sensor models discussed in section 2.3 on page 58 above. The types are MOS: Metal Oxide Semiconductor, and ECM: Electrochemical. Form abbreviations are PTH: Pin Through Hole, a legged electronic component. These tend to be larger than SMD: Surface Mount Devices. Cost: Low means less than ten pounds per unit. Medium means less than a hundred pounds. High means less than a thousand pounds. Performance refers to performance during calibration (next to a reference) in the field. "Good" means	
2.2	achieving a <i>PCC</i> of greater than 0.9	69 82
3.1	Summary of different sensors used in the instrument, with informa- tion from the manufacturer's datasheets. The Sensitivity column for the gas sensors lists species that are meant to have a significant, detectable interaction with the sensor, beyond the target parameter.	90
3.2	Performance statistics for analyses of all three sections. The "typ- ical" values are averages of those statistics across the CAFPOLD timeseries and all four of the sensors in the Long Term Experiment. The SD (standard deviation) values given are the squared sum of the residuals from the timeseries average and the inter-sensor av- erage. The relevant subfigures from figure 3.26 is shown by the bracketed letters	128

3.3	Table of results from the analysis of different equation performance. The SD ratings are the combined residuals from the timeseries av- erage and the average of the sensors a, b, c and d. The linear and	
	multiplicative equations (marked with a $*$ ) for NO <sub>2</sub> sensor b have been omitted from the statistics here. The references to the left of each block of figures is to the graphs in figure 3.27, and the equation numbers are listed next to the equation heading.	130
3.4	Summary of sensor performance statistics of all four sensors, for both $NO_2$ and $O_3$ . The colour grading indicates relative quality, with the darkest green indicating the best performance figure rela- tive to the other sensors, and darker yellow means a higher standard deviation for the performance statistics, and hence more inconsis- tant predictions	194
3.5	Linear regression showing the day zero and change per day statis- tics for RSE, $FE$ and $PCC$ from the CAFPOLD analysis of the LTE predictions using the Inverse equation for NO <sub>2</sub> and O <sub>3</sub> . The uncertainties listed are the 95% confidence interval for the linear	104
3.6	regression	134
37	$RSE$ below $20 \mu g \mathrm{m}^{-3}$	139
0.1	month scale.	140
3.8	Results table for the tested CAFPOLD analyses for $O_3$ on a one month scale. The column labels are identical for table 3.7 on page 140.	140
3.9	Results table for the tested CAFPOLD analyses for $O_3$ on a five	1 4 1
		141

	٠	٠
XXV	1	1

3.10	Table of expected performance of sensors over one month based on the Long Term Experiment dataset, obtained by averaging perfor- mance for the four individual sensors, alongside the performance achieved through the two best tests. 95% confidence intervals are given.	. 144
4.1	Table of statistics from the CAFPOLD analyses of different section- ing techniques, giving the relative difference between the sectioning technique and the baseline. The statistic is obtained by averaging this difference over the whole timeseries. The error incorporates the uncertainty due to averaging between sensors and over time. For all metrics, lower is better except for $PCC$ , where the best results are those closest to 1	157
4.2	Table of performance statistics for each of the sensors using the different fitting methodologies. The statistics are calculated over the whole dataset produced by the above analysis, shown in figures 4.5 and 4.6. The statistics in the transitional fit section are coloured according to whether the transitional fit beat both other fit methods (blue), performed better than the average between the two (green), were worse than average (yellow) or were the worst option (red). The results from the substandard OXc and OXd sensors are highlighted in red in the leftmost column.	. 161
4.3	Relative increase in performance between transitional fit and ei- ther the worst performing of the start and end period calibration for the sensor in question, the best of them, or the average of the performance metrics of the two calibration periods. A positive fig- ure indicates that the transitional fit improved performance. 95% uncertainties are computed from the standard deviation of the dif- ference between the two dataseries. The figures here are averaged between the individual sensor results, although the NO <sub>2</sub> c and d sensors were not used in this average, partially accounting for the higher uncertainty of NO <sub>2</sub> statistics.	. 165
5.1	Table of performance statistics for the passive and active airflow in- struments compared with the reference over the entire experimental	
5.2	Table of statistics for the exponential voltage decay curve following power interruption for the four sensors. Constants are defined in	. 181
53	equation 5.3	. 188 192
0.0	rubic of i CC mean and percentine - between bensor pans	· 104

	٠	٠	٠
XXV	1	1	1

6.1	Table of average correlation of the sensor predictions between each other, divided into two categories: sensor a vs sensor b from the same instrument, and the sensors compared with different instru- ments. Error ranges are 95% confidence interval. The graph also has statistics for how well on average the instruments correlate with
6.2	the AURN reference, using hour-averaged data from the Zephyr prototype instruments
C.1	Comparison between random forest ("Forest") and linear regression ("Regress") performance, in the form of averaged statistics over one month from calibration. The linear regression used the "inverse" equations detailed in section 3.4.1
C.2	As above, but over five months

## Abbreviations

ADC	Analog Digital Converter
AFE	Analog Front End
BBCEAS	Broad Band Cavity Enhanced Absorption Spectroscopy
CAFPOLD	Combined Analysis For Predictions Of Long Datasets
$\mathbf{CFD}$	Computational Fluid Dynamics
CMOS	Ceramic Metal Oxide Semiconductor
DC	Direct Current
$\mathbf{FE}$	Fractional Error
FTIR	Fourier Transform Infra- $\mathbf{R}$ ed spectroscopy
GPRS	General Packet Radio Service
GPS	Global Positioning System
$\mathbf{GSM}$	Global System for Mobile communications $\mathbf{G}$
I/O	Input Output
IR	Infra $\mathbf{R}$ ed
LTE	Long Term Experiment
MCU	Micro Controller Unit
MOS	Metal Oxide Semiconductor
OEM	Original Equipment Manufacturer
PCC	Pearson Correlation Coefficient
PCB	Printed Circuit Board
PTFE	$\mathbf{P}$ oly $\mathbf{T}$ etra $\mathbf{F}$ luoro $\mathbf{E}$ thylene
$\mathbf{RSE}$	$\mathbf{R}$ esidual $\mathbf{S}$ tandard $\mathbf{E}$ rror
SD card	Secure Digital card
SOGS	$\mathbf{S}$ mall $\mathbf{O}$ pen $\mathbf{G}$ eneral $\mathbf{S}$ ensor
$\mathbf{USB}$	Universal Serial Bus
$\mathbf{UV}$	Ultra Violet
VOC	Volatile Organic Compound

Hope is our duty...

### Chapter 1

### Introduction

#### 1.1 Overview

This chapter provides context for the scientific work given in this thesis. In section 1.2 a description of the structure and dynamics of the Earth's atmosphere leads us down into the troposphere. The chemistry of the troposphere, pollutants within it and their relation to human health are discussed in sections 1.2 to 1.3.1, which gives some grounding to the decision to focus on the measurement of NO<sub>2</sub> and O<sub>3</sub>. Finally, an overview of the various instruments available for investigating these chemicals provides a motivation to develop and refine novel small sensors using Metal Oxide Semiconductor technology.

#### 1.2 The Human Atmosphere

The atmosphere of our planet is not a homogeneous sphere of gas. It has a layered structure with considerable variation in composition and temperature. Gravity is responsible for the generally decreasing pressure of the atmosphere with increasing altitude and is a driving force in the movement of gas throughout it. A larger influence on the structure and chemical nature of the atmosphere is solar radiation[1], which is both absorbed throughout the atmosphere resulting in changes in temperature, and is a reagent in photochemical reactions. The interaction between these mechanisms divides the atmosphere into layers, and the most common way to delineate these layers is on the basis of vertical temperature profile[2, p. 28] see figure 1.1



FIGURE 1.1: Diagram of the changes in Earth's atmospheric temperature (bottom axis, solid line) and pressure/density (top axis, dashed lines.) with height. Four atmospheric layers are labelled here, alongside the temperature inversions that define the different layers. The Exosphere starts and extends well beyond the vertical scale of this diagram. Pressure and density both decrease with height, but not quite in the exponential way one would expect mathematically, based on the hydrostatic equation. The different layers are picked out with an orange tint. Adapted from [2, fig 2.6]

#### 1.2.1 Atmospheric structure

The theoretical definition of the upper limit of the Earth's atmosphere is the altitude at which the dynamics of the gases within start to be dominated by the solar wind, rather than Earth's gravity and magnetic field[1]. This is the outer limit of the exosphere, the highest layer of the atmosphere, and is not a solid border but a gradual and blurry transition to interplanetary space. Much of the gas in the exosphere is molecular hydrogen, and the ultraviolet light reflected off this gas forms the geocorona, seen by spacecraft looking back toward Earth with ultraviolet cameras as a glowing shroud around our planet. The approximate extent of this corona is a colloquial definition for the maximum height of the exosphere, some 100,000 km above the Earth's surface[3], although the theoretical boundary is at roughly twice this altitude.

The lowest part of the exosphere, the exobase, is the point at which the Earth's atmosphere becomes dilute enough that it cannot be said to meaningfully have a pressure. Here collisions are so infrequent that the ideal gas equation does not apply. Depending on solar activity, this physical transition is between 500 km and 1000 km from the Earth's surface[4].

Beneath the exosphere is the thermosphere. Density in this region is so low that the mean distance a molecule travels before colliding with another can be several kilometers, and the constituents are simple molecular gases like  $N_2$  and  $O_2$ , and atomic oxygen. These gases are stratified in the thermosphere according to their molecular masses[2, p. 30]. Calculated from molecular kinetic energy, temperature here gradually increases from the most frigid cold found naturally in the Earth environment to over a thousand degrees, but the scarcity of particles means heat is conducted extremely slowly to solid objects like spacecraft, and instead radiation is the dominant means of heat exchange[4].

One consequence of the low density is a lack of chemical reactivity in this part of the atmosphere, and molecules excited by interactions with cosmic rays cannot kinetically dissipate their energy and instead must radiate it as photons[4]. The colourful displays of the aurora thus occur in this layer of the atmosphere. For the purposes of spaceflight, the very low density of the thermosphere means it can be considered as "space", as spacecraft are capable of flying through it without atmospheric drag causing them to dissipate so much of their velocity that they fall back to Earth after a few orbits. It encompasses a significant volume, extending down as far as the mesopause, roughly 85-90 km above the surface.

The mesosphere is the next layer down, and temperature increases with decreasing altitude. This is because the higher gas densities absorb more solar radiation to counteract the dissipation of thermal kinetic energy through radiative emission. Within the mesosphere turbulent mixing ensures that gas concentrations are roughly uniform, and this is the highest location in the atmosphere where  $H_2O$ and  $CO_2$  can be found[1].

The presence of water vapour means the mesosphere is capable of forming (very high, cold) clouds in certain circumstances. The air density is high enough that spacecraft cannot remain in orbit and meteorites begin to burn up within this layer. In the mesosphere the atmosphere starts to act like an ideal gas, and is capable of turbulence on a very large scale, with gravity waves the size of continents. The energy driving the dynamics of the mesosphere is mostly transmitted upward from the denser atmospheric regions [4, p. 30].

The mesosphere transitions into the stratosphere at an altitude of roughly 50 km. The stratosphere is defined by temperature that increases with altitude. This is because the stratosphere contains the ozone layer, which absorbs ultraviolet radiation from sunlight. Where such wavelengths are most intense at the top of the layer, temperatures can approach zero degrees Celsius, but the dwindling UV light further down makes the lowest parts of the stratosphere very cold.

Because of the temperature profile of the stratosphere, with the hottest air already at the top, it is very dynamically stable and doesn't experience the weather that marks the troposphere, although it contains low concentrations of water vapour and occasionally can form clouds. The troposphere exists at altitudes of around 15 km at the equator, but tapers off to 10 km and below at higher latitudes. Because it is heated by the surface, the air lower down in the troposphere is generally warmer than at the tropopause, and consequently convection occurs within it, particularly during daytime. This movement is the driving force behind Earth's weather. Most of the water vapour in the atmosphere exists in this layer, which consequently is the home of precipitation and almost all cloud.

Closest to the surface is the boundary layer, where the motion of the air is directly affected by the geometry of the terrain. The effects of surface features that channel and sculpt the air currents propagate through the boundary layer with a timescale of an hour, or less[5, p. 2]. The rest of the free atmosphere in the troposphere is also affected by the surface, but the timescale can be much longer. The boundary layer's border can be a few hundred meters up on clear, calm days, but generally it lies about a thousand meters from the surface, and it breaks apart and reforms over the day-night cycle.

It is the troposphere can be most easily and directly affected by human activity, and that comprises the habitat for almost all of us. The quality of the air in the troposphere is thus a matter of universal concern.

#### **1.3** Airborne pollution

Many negative health effects of poor air quality are well understood, and the history of serious legislative efforts to mitigate the problem spans over a hundred years in the UK[6]. While early laws concentrated on smoke, including the Clean Air Act passed after the Great Smog of 1952[7], improvements in sensing technology and a better understanding of the environment have changed the focus of air pollution research. Metrics like "black smoke", which were defined empirically by
appearance of the filters in simple pollution detectors, were supplanted by metrics based on the constituent species that make up pollution.

"Pollutant" is a broad term applied to several substances which are harmful to human life, and whose presence or elevated concentration in the environment is due to human activity[8]. While many pollutants are harmful when directly ingested or inhaled by humans, some instead affect humans indirectly by influencing the ecological systems necessary to sustain life, and many airborne pollutants have different, even positive effects depending on their location in the atmosphere. Ozone, for example, is a toxic gas if it is produced or transported to ground level and inhaled, but in the lower stratosphere it exists as part of a photochemical system, which along with oxygen, is responsible for absorbing the majority of the sun's incident radiation at ultraviolet wavelengths of less than 290 nm, which otherwise would be harmful if it reached the Earth's surface[9].

Pollutants exist as part of a complex and interwoven system of chemical interactions, with some being emitted into the atmosphere directly (primary air pollutants) and some being generated through chemical reactions in the atmosphere itself (secondary air pollutants). These interactions vary depending on the location and time of day, and mean that different environments (indoor and outdoor, urban and rural) will typically contain different sets of pollutants.

## 1.3.1 Pollutants in the urban environment

Measurement of  $NO_2$  and  $O_3$  in urban environments is the focus of this thesis, for which the following text provides context and justification. This section contains an overview of the types of airborne pollution found in urban environments, elaborates on the ways they directly affect human health, as well as their sinks, sources and interactions. The different types of pollution environment throughout the troposphere are defined, and their characteristics given, in section 1.3.3. Section 1.3.5 on page 22 is a closer examination of the way pollution moves through cities, shaped by their buildings and industry.

Most pollutants can be measured in several ways, and standardized methods of doing so have been developed by both the Department for the Environment and Rural Affairs (DEFRA) in the UK and the Environmental Protection Agency (EPA) in the United States. DEFRA runs the Automatic Urban and Rural Network (AURN)[10], a set of air monitoring stations that ensure compliance against the European Union's Ambient Air Quality directives, which provide most of the statistics in this section. In addition to these, this section makes use of the EPA's equivalent method report[11]. Of all these detection methods, particular attention will be paid in this section to sensor types that are small enough to be portable, as well as selective.

#### 1.3.1.1 Particulate matter

Particulates consist of a variety of solid and liquid chemicals, suspended as an aerosol in the atmosphere. For examining health effects they are classified based on diameter, not composition, as this determines both their movement and lifetime in the atmosphere and how easily they can penetrate into people's lungs [12, p.58]. Particulate matter is defined in two categories for the purposes of legislation,  $PM_{2.5}$  and  $PM_{10}$ , being those particles with diameters that let them pass through a size-selective inlet with a cutoff at 2.5 µm and 10 µm respectively[13]. Such inlets use inertia to filter the different classes of particle and are required for non-optical measurement instruments. Beyond their potential for cardiovascular and respiratory system injury through long-term inhalation[14], particulate matter also influences the broader biosphere in important ways and anthropogenic particulates have been linked to suppressing rainfall by 10-25% in some areas[15].

The dynamics of particulate formation and growth are complex and diverse. Particulates might be emitted directly as soot from combustion reactions, dust from abrasive processes or sea spray, or produced as secondary particles. They grow as they accrete soluble particles from the atmosphere. The interior of waterlogged particles allows for aqueous chemical reactions to take place on a large scale in the polluted atmosphere[16].

The average urban background concentration of particulate matter in the UK was  $14.7 \,\mu g \,m^{-3}$  for  $PM_{10}$  and  $10 \,\mu g \,m^{-3}$  for  $PM_{2.5}$  in 2018. Roadside levels have been around  $3 \,\mu g \,m^{-3}$  higher for  $PM_{10}$  and  $1 \,\mu g \,m^{-3}$  higher for  $PM_{2.5}$  for the last ten years[17]. Peaks of above  $35 \,\mu g \,m^{-3}$  are useful for detecting short term high pollution events, and dust storms and bush fires are capable of increasing  $PM_{10}$  concentrations to over  $300 \,\mu g \,m^{-3}$ [18] for several hours.



FIGURE 1.2: A gravimetric sensor. Air enters through **A**, a size-selective inlet which rejects particles above a certain mass, and is drawn through the filter **B** by the pump **C** at a constant rate.

The in-stack particulate filtration technique is the Federal Reference Method for counting particulate matter[11]. Also known as "Gravimetric" sensors, these devices have a long history and are accepted as an equivalent reference by DEFRA[19, p. 128]. Figure 1.2 shows the operation of

a gravimetric sensor, which loads particulate matter into a filter. After the sampling period is over, the filters are then extracted from the instrument, dried out and weighed. The mass increase of the filter during the experiment, combined with the volume of air pumped through it, can be used to calculate the mean density of particles in the atmosphere.

Sensors of this type are reliable due to their mechanical simplicity, and are the most commonly deployed in the USA. However extracting and weighing the filter must be done by hand, a major disadvantage. The sampling period depends on how attentive the operator is, and day-long periods are not uncommon. A related method was commonly used some fifty years ago in the UK when much of the particulate matter pollution was in the form of black soot from domestic fires. The "black smoke" sensors had a similar principle of operation to a gravimetric sensor, except that the filter was scanned with a reflectometer instead of being weighed, with a darker stain implying a higher concentration of soot in the atmosphere.



FIGURE 1.3: Schematic of a Beta Attenuation Monitor. Air enters through the size selective inlet A, and is pulled through a porous ribbon by the pump
B. The ribbon drawn from reel C onto
F, across the airstream at a steady speed, and after exposure it passes under a beam of beta radiation from a radioactive source D. The more particle mass is stuck in the ribbon, the more attenuated the radiation will be when it reaches the detector E.

The "black smoke index" from this method was not directly comparable to  $PM_{10}$  or  $PM_{2.5}$ , and no reliable relationship could be found between the colour of the filter and the amount of soot in the atmosphere, because different particle compositions could have different albedos and mislead the reflectometer [19, p. 134]. Despite these shortcomings, long term experiments using this method have been carried out around the world and a few such sensors are still in operation.

The AURN still makes use of gravimetric sensors, but for a fully automated network, methods that are less manually intensive

are preferable. BAM, TEOM and the more specific FDMS instruments are also used in the network as equivalent sensors[20] for  $PM_{2.5}$  and  $PM_{10}$ .

Figure 1.3 shows the operation of a Beta Attenuation Monitor (BAM). BAM instruments are the second most common means of detecting particulate matter among EPA weather stations in the USA and UK, although they are the most popular in mainland Europe[19, p.131]. They are more precise than filter-based sensors, and can give continuous real-time readings at hourly rather than daily intervals, although the presence of a radioactive source complicates their deployment.



FIGURE 1.4: Schematic of a Tapered Element Oscillating Microbalance. Air enters through the size selective inlet **A**, before splitting in two at **B**, to create a bypass that is used to regulate the amount of air passing through the sensor element, a tapered glass cylinder **C**. One end of this cylinder is fixed to the instrument, but the other is free to vibrate at its natural frequency. This frequency is continually measured electronically **D**, and as particles accrete onto the cylinder, their mass affects this frequency, like a singing wineglass changing its note as your pour in more liquid. The valves  $\mathbf{E}$  control the amount of air flowing through the sensor and by-

pass line to the exhaust pump  $\mathbf{F}$ .

Figure 1.4 shows a Tapered Element Oscillating Microbalance (TEOM) sensor. By using an inlet to only permit particles of a desired diameter or smaller, measuring the mass deposited on the filament over time and controlling the air flow rate through the sensor, the concentration of particulate matter can be determined.

The Filter Dynamics Measurement System (FDMS) used by the AURN network contains a TEOM-based sensor and additional filters that dry out a portion of the inlet air, allowing the fraction of volatile to non-volatile particulate matter to be determined[21].

TEOM sensors use proprietary technology from Thermo-Scientific and can only be found in instruments made by that com-

pany. The University of Leicester AURN station makes use of an FDMS to give hourly PM readings.

There are several methods that measure particulate matter using its interaction with dust. The simplest way is to simply measure air opacity, using a light source and detector. Transmissometers can have a very long open path length, while nephalometers such as the inexpensive Shinyei PPD42[22] are compact and use a photodiode and LED for the atmosphere within a small housing (see figure 1.5. These methods give a qualitative indication of such pollution, although like black soot sensors there is no direct relation to  $PM_{2.5}$  or  $PM_{10}$ . A PPD42 sensor was used as part of the instrument developed over the course of this thesis, although the results are not relevant to the scope of this work.



FIGURE 1.5: Schematic of a Shinyei
PPD42 nephalometer. Air beneath the housing A is heated by the resistive element B, drawing air up through the sensor. An infrared LED C shines into this air, and the diffuse reflected glow from particles within the rising stream D will be detected by the photodiode E, which is obscured from the LED.

Light interacting with roughly spherical particles undergoes Mie scattering, producing an angular spectrum related to the size of the particle. Optical dust sensors contain an array of photosensors and a laser that can produce and capture a Mie scattering spectrum, and process it to return a mass spectrum of particulates. These instruments are small enough to be handheld and seem an attractive solution for small sensors, except that once again their readings depend on assumptions about the composition of particulates that are difficult to verify[19, p.133].

These are the principal current and historical methods for in-situ measurement of particulate matter for air quality monitoring purposes. More advanced instruments such as aerosol mass spectrometers and cloud condensation nuclei counters are able to provide more detailed mass spectra for atmospheric aerosols, but they are complex and expensive enough to be overkill for day-to-day air quality monitoring purposes. They find a use among scientists developing new techniques as well as for PM source attribution[23].

#### 1.3.1.2 Sulphur compounds

Sulphur pollution is produced primarily from fossil fuel burning in power generation[13]. Unrefined coal and oil contain up to 5% sulphur compounds by volume [24, p.396] and these are converted directly into sulphur oxides when the fuel is burned. Refined fuels frequently have had a large proportion of the sulphur removed, although shipping, power generation and domestic heating frequently do not use such fuels and consequently are a major anthropogenic source of  $SO_2[13]$ . Metal smelting is also a significant source in industrialized areas, but thanks to changes in fuel treatment  $SO_2$  does not generally exist in dangerous concentrations in European cities.

That the sulphur dioxide is removed from some fossil fuels but not others makes it a potential tracer for its specific sources of pollution, but only in concert with measurements of other fossil fuel byproducts.



FIGURE 1.6: Schematic of a UV fluorescence sensor. Air  $\mathbf{A}$  is pumped through the sensor and exposed to light pulses from lamp  $\mathbf{B}$ , at a frequency capable of exciting SO<sub>2</sub> molecules. The photons these re-emit pass through the optical filter  $\mathbf{C}$  and are detected by the photodiode  $\mathbf{D}$ .

a gas such as  $SO_2$ , which is by far the most common sulphur compound in the troposphere and causes acute health issues if inhaled[13], and condensed as a secondary particulate, in the form of sulphite (for example (NH<sub>4</sub>)<sub>2</sub> SO<sub>4</sub>). Sulphur in this form can significantly affect the climate through scattering solar radiation[25], in addition to potentially causing injury as a particulate

Sulphur can exist in the atmosphere both as

#### aerosol[14].

UV florescence monitors are the DEFRA standard for measuring atmospheric  $SO_2[26]$ . A diagram showing the mechanism of this sensors action is given in figure 1.6. The lamp within such sensors can be pulsed or continuous and typically emits wavelengths of 214 nm. The SO<sub>2</sub> re-emits longer UV wavelengths (typically between 300 nm to 390 nm), although the band-pass optical filter is necessary to remove emissions from other fluorescing gases, such as NO[27]. A higher intensity of florescence corresponds to higher SO<sub>2</sub> concentration.

In the UK, SO<sub>2</sub> levels have remained low for decades. DEFRA counts "moderate" SO<sub>2</sub> levels as being higher than  $267 \,\mu g \,m^{-3}$ , and between 2010 and 2018 there has not been a single year when the average urban measurement site was experienced more than a day over the year when SO<sub>2</sub> concentrations exceeded this threshold[17]. Understandably given the sources of SO<sub>2</sub> in the troposphere, rural sites have even less of a problem.

#### 1.3.1.3 Volatile organic compounds



FIGURE 1.7: Schematic of a 2D Gas Chromatograph. Air is pumped into the sensor and split into two streams at A. The airstreams pass through valves which controls the pressure in the 'columns' C and D. The two columns contain different granular material, through which different gas species will percolate at different speeds. At the end of the columns is a Flame Ionization Detector (FID) E, which ignites the gas output (the 'eluate') in a hydrogen flame and detects the intensity of the ionized combustion products using electrostatic plates **F**.

The term "Volatile Organic Compound" (VOC) encompasses a range of hydrocarbons that have a high enough vapour pressure that they will remain in the atmosphere in normal conditions. EU Law states that any hydrocarbon with a boiling point at standard atmosphere of less than 250 centigrade may be a VOC[28].Humans contribute VOCs to the atmosphere principally through the use of fossil fuels, both from incomplete burning in engines and fuel evaporation. VOCs consist of many different long-chain hydrocarbons, aromatic compounds and formaldehyde [13]. Plants are a major natural source of VOCs, and isoprene  $(CH_2=C(CH_3)-CH=CH_2)$  and other ter-

pene compounds dominate their emissions<sup>[29]</sup>. VOCs have enough in common chemically, but enough diversity in their structure and sources, that they can be measured individually using gas chromatographs to allow for quite advanced source attribution<sup>[30]</sup> - see figure 1.7. VOCs can cause respiratory disease by themselves, with the aromatic compounds often produced as a byproduct of combustion being linked to cancer and other morbidities[8]. Aside from their direct health effects, VOCs play an active role in tropospheric chemistry, chiefly in the production of ozone[31] and secondary organic aerosols, which can occasionally dominate aerosol concentration[32]. Methane is also an important greenhouse gas with an atmospheric lifetime of around 8 years[33].

The importance of VOCs, particularly in their production of secondary pollutants[16], make them a worthy candidate for study. However, while broad VOC sensors are available, the diversity of chemical structure among VOCs means that attributing measured pollution to a specific source is difficult if only total VOC information is available. Instruments capable of distinguishing VOC species, such as two dimensional gas chromatographs, mass spectrometers and Fourier Transform Infrared Spectrometers (FTIR), are all bulky and expensive devices. While they have given important results as part of measurement campaigns in the field, their size and cost precludes their use in small sensor networks.

#### 1.3.1.4 Carbon dioxide

The natural consequence of burning any carbonaceous fuel in air, carbon dioxide is the most significant human-produced pollutant that affects the Earth's climate[34, p.678]. Carbon is removed from the atmosphere through several processes such as uptake in the biosphere and dissolving in the oceans[35], which complicates estimation of a typical atmospheric lifetime.  $CO_2$  molecules can last in the atmosphere on a scale of decades, but perturbations of overall  $CO_2$  concentration can persist for hundreds of years, and the climate effects resulting from this can last for thousands[36].

Unlike many other pollutants it is not directly toxic to humans unless its concentration reaches upwards of  $100,000 \text{ mg m}^{-3}[37]$  - around 5% concentration by volume. Global mean ground level  $CO_2$  concentrations have every year for decades, and as of April 2018 surpassed 748 mg m<sup>-3</sup>[38]. Note this is of order four orders of magnitude more abundant than pollutants like methane and nitrogen oxides.



FIGURE 1.8: A Nondispersive Infrared  $CO_2$  sensor. A lamp produces infrared light **B** of a specific wavelength which is absorbed by  $CO_2$  molecules in the air **A**, causing attenuation that can be detected by the properly calibrated photodiode **C**.

Because of its long lifetime,  $CO_2$  is well mixed throughout the troposphere. Compared to the substantial variability of pollutants like NO<sub>2</sub>, mean CO<sub>2</sub> concentration at the tropopause varies annually by around 3%[39], and it decreases with altitude through the troposphere by a similar amount[40].

Consequently, despite  $CO_2$  emission being critically important for the study of climate change, and being easy to measure with small Nondispersive Infrared (NDIR) sensors (see figure 1.8), it is not a very precise or useful tracer of specific pollution sources.

#### 1.3.1.5 Heavy metals

Another broad category containing metals such as lead, nickel, cadmium, mercury and arsenic. The high toxicity of these substances has driven strong emissions regulation; for example, after lead additives were banned from petrol in the UK in 1999, emissions of airborne lead compounds fell by 98%[13]. These pollutants can also poison groundwater and soil. Industrial processes and fossil fuel burning are now the largest sources of heavy elements in the UK atmosphere.

#### 1.3.1.6 Nitrogen oxides and ozone

In the urban atmosphere two nitrogen oxides,  $NO_2$  and NO (known together as NOx), are important pollutants. Together with ozone  $(O_3)$ , which is harmful in

its own right, they exist in a closely linked null cycle reaction system (see section 1.4 for information). Two thirds of the nitrogen oxides present in the atmosphere are a result of fossil fuel burning [41], and in the UK road transport is the largest single source, accounting for a third of all such emissions[13]. The remainder is evolved from biomass and produced by lightning strikes[41].

Around 85% of ozone in the atmosphere as a whole is produced from photolysis of oxygen molecules with UV light in the stratosphere, a process obviously not tied to human activity, but only around 10-15% of the ozone that reaches the troposphere is transported from above. Most of the ozone in the lower atmosphere ozone is a secondary pollutant, produced in photochemical reactions that require both NOx and hydrocarbons. While VOCs and NOx exist naturally to a degree, human industry can supply both of them and this mechanism is responsible for an estimated 39% of ozone in the troposphere[42].

The reactions between NOx and  $O_3$  occur on a timescale of a few minutes during the daytime, but sinks of NOx out of the atmosphere can take several hours (6 hours in the summer, 18-24 in winter)[43], short enough that NOx does not become widely dispersed from its source, but long enough that the pollutant gas can diffuse from a source and still be detectable, allowing good spatial and time resolutions for pollution mapping. In the absence of any interfering factors, these gases have concentrations predicted by the Leighton relationship[44], but in practise each species has its own distinct sinks and sources related to other parts of atmospheric chemistry.

 $NO_2$  and  $O_3$  pollution has well documented adverse health effects [45][13].  $NO_2$  in particular causes acidification and eutrophication of soil, groundwater and surface waters and can decrease ecosystem vitality and biodiversity, and cause groundwater pollution [46][47]. Elevated tropospheric ozone can cause respiratory disorders as well as reduced crop yields[48]. Ozone levels are higher on average in rural environments; moreover ozone production often is highest in rural areas, downstream of urban sources (see section 1.3.3.2 on page 19).

The importance of NOx as both a health risk and a tracer for vehicular and industrial pollution has led to a variety of measurement techniques being developed and deployed. Satellite retrievals using the OMI instrument [49] have been developed and are fairly mature, and chemiluminescence sensors are the government standard for measuring NOx in both the EU and USA. Smaller and cheaper sensors for measuring nitrogen oxides have transitioned from their original purpose as workplace safety warning devices and are beginning to be used in air pollution monitoring applications[50]. More detail on measurement techniques for nitrogen oxides is given in section 1.5 on page 31

Average NO<sub>2</sub> concentrations at urban background sites have steadily reduced over the last fifteen years and were 20.1 µg m<sup>-3</sup> as of 2018[17]. At rural sites over the same period concentrations have been within a microgram of 7.3 µg m<sup>-3</sup>. Roadside concentrations in 2018 were much higher at 32.8 µg m<sup>-3</sup>. This variation in concentration underscores how useful NO<sub>2</sub> can be in identifying pollution hotspots. Average O<sub>3</sub> concentrations in 2018 were 62.8 µg m<sup>-3</sup> at urban and 73.2 µg m<sup>-3</sup> at rural measurement sites, higher thanks to the scarcity of NO in this environment[17].

## **1.3.2** Pollutant focus in this work

The combination of potential for adverse health effects, usefulness as a traffic emissions indicator for other pollutants, potential for more advanced chemical investigations, multitude of relevant measurement techniques and, most importantly here, the opportunity for genuine frontline research into new sensors make nitrogen oxides and highly attractive as the topic of this research.

## **1.3.3** Classifying environments

Airborne pollution environments can be divided into three main categories - Indoor, Rural ambient and Urban ambient pollution. Each represents its own unique characteristics and issues.

#### 1.3.3.1 Indoor

Indoor pollution is a significant issue for global health, particularly in developing countries. Around half the human population rely on burning biomass for domestic energy, and while it is hard to estimate precisely, the effects of the particulate aerosols released by these fuels might be up to 4% of the world's total disease burden [51][52] through respiratory and cardiac illness and cancer. Social and technical efforts to reduce this toll are underway - since 1980 the Chinese government, for example, has provided simple stoves to rural families to replace their open hearths and allow smoke to safely leave the living area[51].

However, modern buildings in affluent areas have their own set of issues. The increasing emphasis placed on energy efficiency has led to building designs becoming more airtight, and this can cause issues with air quality [53], as any pollutants emitted inside the building are slower to percolate out into the larger environment. Dust particles, volatile organic compounds (VOCs) given off by household chemicals and paint, and combustion products from cooking and domestic heating are all potentially harmful. The chemical environment is very different inside to outside, even though the latter may significantly influence the former [54].

The existence of "Sick Building Syndrome", which produces a variety of symptoms such as fatigue and poor concentration in office workers, has been acknowledged since the 1970s and the term was first coined by the WHO in 1986[55]. The "illness" is thought to be caused by poor indoor air quality (possibly caused by molds in the walls and air ducts)[56] and steps taken to mitigate it include better design of air conditioning systems and making efforts to improve air circulation. Total VOC concentrations and particulate matter counts are both indicators of the freshness of air inside a building, although sick building syndrome is more usefully correlated with specific VOCs and mold or fungal spores[57][56].

#### 1.3.3.2 Rural Ambient

Rural areas have very different industrial output compared to cities, and in much of the world the domestic sources covered in the preceding section provide the majority of the pollutants inhaled by people who live in the countryside. The outdoor rural environment has a lower industrial pollutant load, and sparser rural populations limit total human exposure from individual sources[13]. However, agricultural pollution can be significant; in the UK pesticide and nitrate fertilizer use contributes to water pollution, and ammonia evaporation from livestock farming is responsible for 80% of the ammonia in the atmosphere. Both of these pollutants have negative effects on human health[58].

One of the most important rural pollutants is ozone[59]. Ozone production rates depend non-linearly on the concentration of nitrogen oxides and volatile organic compounds (VOCs) in the troposphere during the daytime. Freshly emitted vehicular nitric oxide reacts with ozone to form NO<sub>2</sub>, a process that often suppresses ozone levels in urban areas. The processes that remove NOx from the atmosphere become more efficient at the higher concentrations of these gases typical in cities, reducing the opportunity for NOx molecules to engage in the catalytic cycles that produce O<sub>3</sub> and capping O<sub>3</sub> production rates. In rural areas with more moderate NOx levels and low nitric oxide, ozone production can be far higher and the presence of VOCs due to biomass burning[60] or extraction industries[61] can cause large spikes in ozone concentration.

The mechanism of ozone production in both urban and rural environments is quite complex, and is discussed further in section 1.4.1 on page 28.

Among the most toxic and carcinogenic of the VOCs found in rural environments are the polycyclic aromatic hydrocarbons (PAH), which are released through biomass burning, and some industrial processes[13] - this is obviously more of an issue in some regions than others[59].

Plants naturally evolve reactive hydrocarbons like isoprene and terpenes, which can cause respiratory irritation in sensitive individuals. They are a major participant in the production of secondary pollutants in rural environments, and are directly linked to the elevated ozone concentrations in green areas at ground level[62][63]. Isoprene is the most abundant non-methane VOC in the atmosphere, and the largest contributor to Secondary Organic Aerosols (SOAs) globally[64]. SOAs, a form of particulate matter, are seeded by many VOCs after oxidisation through reactions initiated by OH or NO<sub>3</sub> radicals. Human activities in rural environments can contribute to this process, for example through straw burning[64].

#### 1.3.3.3 Urban

With over half of the world's population living in cities since the year 2000[65], managing urban air quality is a serious concern. The World Health Organization has a summary of what are considered major urban airborne pollutants, as well as maximum safe exposure ratings [66], but it is now being recognized by both scientists and legislators that there is no 'safe' level below which many pollutants are benign[67]. The linear correlation between illness and death and exposure in urban settings is empirically supported by a growing body of literature in cities across the world [68] [69] [70] as well as meta-analyses tying pollutants to specific illnesses such as stroke[71], heart attack[72] and chronic lung failure[45].

The sources of air pollution most associated with cities are automobiles and industrial activity[65], with the latter being more pronounced in industrializing countries and the former remaining a problem particularly as prosperity and scale increases. Depending on a city's location and surroundings, long distance transport of gas and particularly aerosols can comprise a significant fraction of a city's exposure. But ground-based measurements [23], a long history of simulations[73][74][75] and airborne spectrographic mapping[76] all indicate that road traffic (and combustion in general) accounts for a large proportion of urban pollution, across a range of pollutants from aerosols to sulphur compounds to nitrogen oxides[77]. Particularly indicative is the fact that even relatively primitive models, if traffic data is the only input, have substantial predictive power in air quality measurements taken at ground level for a variety of pollutant species[78]. Regional sources, both industrial and agricultural, can be significant enough that a purely local focus is insufficient for understanding and mitigating pollution[79] but if anything these cases underscore that a holistic approach to tackling air quality requires reducing pollutant output at the source as well as dispersing its production.

The large scale of the problem, its societal cost and the potential for improvement through intelligent city planning and regulation make urban air quality a high priority for scientists and it will be the focus of this thesis. Every city in the United Kingdom is bound to produce an air quality action plan, evaluating different policy approaches in terms of their effects on urban pollution. Leicester's plan [80] places a justified emphasis on reducing traffic emissions, but specific proposals to this effect require both accurate measurements of pollutants on an urban scale and useful predictions of the consequences of legislative action.

## 1.3.4 Dynamics of the boundary layer

The sources of pollution are only one part of the puzzle in urban environments. The motion of the atmosphere, which transports and disperses pollutants after they've been emitted, determines human exposure to those sources. This movement is most complex and relevant in the boundary layer, the layer in which humans (and most of the sources of pollution harmful to humans) reside.

Compared to the rest of the atmosphere, direct solar warming of the boundary layer is not a significant driver of the air's temperature or motion. Instead, the sun warms the ground, and this heat is transmitted by contact into the air near the surface. Turbulence is a key force throughout the boundary layer, and is highly effective at propagating pollution, heat and air velocity throughout it[5, p. 3]. This effect keeps the boundary layer well-mixed by day, although it settles and becomes stratified at night.

Turbulence in the boundary layer comes from a number of sources - solar heating of patches of ground can cause large thermals to form. Moving air produces wind shears from friction with flat surfaces, and turbulent wakes form downstream from obstacles like trees and buildings. A mixed layer forms in the boundary layer when turbulence has produced a uniform vertical distribution of air velocity, temperature and pollutant contents. Conversely, when the ground is colder than the air (often at night), the boundary layer becomes stable and stratified[5, p.16].

Pollution emitted into the boundary layer from sources forms plumes, the downstream shape of which is governed by the size of turbulent eddies, and turbulence also mixes the gases from elevated plumes down to ground level. Conversely, pollution plumes can be lifted away from the ground if a stable region of the boundary layer forms beneath the pollution plume[5, p.19].

## **1.3.5** The interactions of air in cities

The urban environment has quite complex air dynamics, and there are many considerations for an air quality manager[81], only a few of which are summarised in this section. Man-made structures affect the motion of air in the boundary layer, and so the geometry of cities influences air pollution exposure[82, 19]. While buildings and roads in cities are more commonly shaped by the forces of the local economy, these structures are starting to be designed with features to mitigate

pollution, and urban vegetation is increasingly seen as a solution that is both easy to implement and contributes to livability[83].

Vegetation, in the form of trees and hedges as well as artificial structures like green walls and roofs, is capable of absorbing airborne pollutants. However, it also changes the movement of the air and hence affects the rate at which pollution disperses[84].

Figure 1.9 shows profiles of four common environments involving roads - two with vegetation, two without; two in a street canyon and two involving an open roadway. The four environments are characterized by different dynamics, the most



FIGURE 1.9: Four representations along two axes of road environments, with a cars and pedestrians nearby. Blue arrows indicate wind direction, and the green shaded area represents an example shape of an exhaust plume in those environments. Adapted from [83].

important for pollution mitigation being how likely it is for the road pollution to be channeled away into the higher parts of the atmosphere before it can be inhaled by nearby pedestrians.

• Road, no vegetation: Exhaust gases can easily be breathed by nearby pedestrians, but are not necessarily trapped there.

- Road, vegetation: With either a hedge between the road and pedestrian walkways, the road pollution is either adsorbed or deposited onto the vegetation, or lifted away from the ground as it travels over it, making it less likely to be breathed in by pedestrians provided the air is not too calm.
- Street canyon, no vegetation: Wind blowing across the roofs of buildings in a street canyon causes circular eddies within it. If the weather promotes ventilation, gas produced at ground level can be lifted away from the road and into the atmosphere above.
- Street canyon, vegetation: Contrary to vegetation's beneficial effects in the open road case, tree canopies can interfere with downward or upward air flows, blocking the circulation that is required to disperse pollution away from head-height and the road and potentially increasing exposure.

While these examples are far from exhaustive, they demonstrate that there is no simple solution to this issue for urban planners. Computational Fluid Dynamics (CFD) can be used to examine the flow of pollutant gases and their interaction with buildings and vegetation[84], but a summary of this tool is beyond the scope of this work.

However, as shown in Abhijith et al. [83], generally the closer people are to a road, the more pollution they will be exposed to, that hedges and especially parks can reduce exposure considerably; conversely street canyons with trees produce poor air quality, particularly where there is congestion or collections of idling vehicles.

# 1.4 Atmospheric chemical reactions of nitrogen oxides

Further discussion of the research goals of this work will require some grounding in nitrogen oxide atmospheric chemistry in the troposphere. Material underlying the following discussion can be found in reference [12].

### 1.4.1 Nitrogen oxides and ozone

The gases within an internal combustion engine can reach temperatures of a thousand degrees centigrade, causing the decomposition of molecular oxygen, which then catalyses a reaction between molecular nitrogen and molecular oxygen to produce nitric oxide (NO) as shown in reactions 1.1 and 1.2 [85].

$$O + N_2 \longrightarrow NO + N$$
 (1.1)

$$N + O_2 \longrightarrow NO + O$$
 (1.2)

Postemission, nitric oxide and nitrogen dioxide are bound together by a set of reactions, such that in the troposphere they have similar typical concentration. One set of interactions occur after the nitric oxide produced above has left an engine, when its concentration is still high enough that it can combine into nitrogen dioxide as in reaction 1.3.

$$NO + NO + O_2 \longrightarrow 2 NO_2$$
 (1.3)

Generally in the atmosphere NO concentrations are too low, and hence collisions between NO molecules are too rare, for this reaction to be important. Instead, NO reacts with atmospheric ozone to give NO<sub>2</sub> (reaction 1.4), and the NO<sub>2</sub> can be photolysed to produce an oxygen atom by sunlight (h $\nu$ , reaction 1.5). The oxygen atom will rapidly react with molecular oxygen to regenerate the ozone. This involves an anonymous nearby molecule M, collisions with which remove the additional energy produced when the new chemical bond forms. M is left unscathed after reaction 1.6 [44].

$$NO + O_3 \longrightarrow NO_2 + O_2$$
 (1.4)

$$NO_2 + h\nu \longrightarrow NO + O(^{3}P)$$
 (1.5)

$$O + O_2 + M \longrightarrow O_3 + M$$
 (1.6)

The timescale for these reactions is a few minutes in daylight, although at night these reactions stop and another set occur producing the nitrate radical  $NO_3$ , which will be discussed later. The oxidation of NO to  $NO_2$  and back has such a short timescale that they are referred to collectively as NOx. The proportion of  $NO_2$  to NO within NOx, in the absence of other factors, is given by the Leighton Relationship:

$$\frac{[\mathrm{NO}]}{[\mathrm{NO}_2]} = \frac{\mathrm{J}}{\mathrm{k}[\mathrm{O}_3]} \tag{1.7}$$

Where k is the rate coefficient of the NO +  $O_3 \longrightarrow NO_2 + O_2$  reaction 1.4 at a given temperature, and J is the photolysis rate constant for NO<sub>2</sub>, which depends on solar radiation levels and thus varies with time, weather and location. Square brackets signify concentration. NOx is a useful metric because while the balance NO<sub>2</sub> and NO can shift on a scale of seconds, the total NOx is much longer lived. The relationship is derived from the equilibrium concentrations of the three gases, in a photochemical steady state, given the rates at which they react.

While the interaction described by the Leighton relationship is important, it does not always hold. Additional gases supplied to an environment will make the Leighton relationship a less reliable predictor - for example, freshly emitted NO from car exhausts makes the relationship particularly unreliable in urban areas. One important such unbalancing factor involves volatile organic compounds. VOCs are oxidised by hydroxyl (OH) radicals, which themselves are produced in the troposphere primarily through the photolysis of ozone into molecular oxygen and electronically excited atomic oxygen (the excitation is denoted by the term symbol label <sup>1</sup>D) in reaction 1.8. The excited oxygen has enough energy to initiate a reaction with water (reaction 1.9). The resulting hydroxyl radicals have a lifetime of a few seconds due to their very high reactivity, and their propensity for abstracting hydrogen atoms from other compounds (particularly VOCs) make them a major initiator of chemical reactions in the lower atmosphere during the daytime.

$$O_3 + h\nu \longrightarrow O(^1D) + O_2$$
 (1.8)

$$O(^{1}D) + H_{2}O \longrightarrow 2 OH$$
 (1.9)

Hydroxyl radicals readily react with VOCs (denoted RH), robbing them of a hydrogen atom and producing an alkyl radical R that immediately reacts with oxygen, giving an organic peroxy radical.

$$OH + RH \longrightarrow H_2O + R$$
 (1.10)

$$R + O_2 + M \longrightarrow RO_2 + M \tag{1.11}$$

Hydroxyl radicals also react with carbon monoxide (CO) but with a different mechanism, which produces hydroperoxy (HO<sub>2</sub>) radicals. While not a VOC per se, carbon monoxide is also found in car exhausts.

$$OH + CO \longrightarrow H + CO_2$$
 (1.12)

$$H + O_2 + M \longrightarrow HO_2 + M$$
 (1.13)

The peroxy radicals produced in the previous two pairs of reactions react with nitric oxide to produce nitrogen dioxide, as well as regenerating the peroxy radicals and producing secondary VOCs (aldehydes) in the case of the organic peroxy reaction.

$$\operatorname{RO}_2 + \operatorname{NO} \longrightarrow \operatorname{RO} + \operatorname{NO}_2$$
 (1.14)

$$RO + O_2 \longrightarrow R'CHO + HO_2$$
 (1.15)

$$HO_2 + NO \longrightarrow OH + NO_2$$
 (1.16)

Reaction 1.16 closes the cycle by converting  $HO_2$  to OH, such that the hydroxyl radicals can then proceed to react with another VOC. Because the peroxy and hydroxyl radicals are regenerated, whenever there are VOCs available NO will be oxidised to NO<sub>2</sub>. Second, photolysis of the NO<sub>2</sub> will lead to the production of tropospheric ozone (reaction 1.5).

Since ozone is consumed in the process of generating hydroxyl radicals in the first place, and the two nitrogen oxides are so closely linked together, the first important conclusion is that ozone production during the day is a process that has two requirements - the presence of VOCs, and the presence of NOx. If there is not sufficient NOx, reaction 1.15 can proceed to:

$$HO_2 + O_3 \longrightarrow OH + 2O_2 \tag{1.17}$$

which consumes  $O_3$  in very clean regions of the atmosphere. Peroxy radicals can also react with each other to produce aldehydes and peroxides, taking them (and by implication the ozone that produced them) out of circulation. And without sufficient VOCs, there is nothing to replace ozone in the nitrogen oxide system and it will thus not build up in the atmosphere.

For this reason, the level of ozone during the day can be sensitive to either NOx or VOCs, depending on their relative abundance, as shown in figure 1.10. In Beijing, VOCs and aerosols from multiple sources dominate, whereas NOx from traffic pollution dominates the atmosphere in Paris.[12, p.3] Thus in the former, ozone



FIGURE 1.10: Isopleths giving the net rate of ozone production (in part-perbillion per hour) as a function of VOC (part-per-billion carbon) and NOx (partper-billion) concentrations, for mean summer daytime meteorology and clear skies. The curved lines represent production rates of 1, 2.5, 5, 10, 15, 20, and 30 ppb/h. Image credit[86].

concentration tends to follow NOx (ie. Beijing is NOx limited), and in the latter it instead follows VOC levels (VOC limited).

An important sink for both peroxy radicals and  $NO_2$  is conversion to highly soluble nitric acid (HNO<sub>3</sub>)

$$OH + NO_2 + M \longrightarrow HNO_3 + M$$
 (1.18)

At night, when hydroxyl radicals are no longer produced and nitrogen dioxide is no longer photolysed, ozone and nitrogen oxides react to produce the nitrite radical NO<sub>3</sub>:

$$NO + O_3 \longrightarrow NO_2 + O_2$$
 (1.19)

$$NO_2 + O_3 \longrightarrow NO_3 + O_2$$
 (1.20)

The resulting nitrate radical acts as the major oxidizing species at night, but it is rapidly photolysed during the day. It exists as part of a rapid equilibrium system with nitrogen dioxide to form the reservoir species dinitrogen pentoxide:

$$NO_2 + NO_3 + M \longleftrightarrow N_2O_5 + M$$
 (1.21)

The dinitrogen pentoxide can react with water to produce nitric acid  $HNO_3$ , the deposition of which is the main night-time sink of NOx from the atmosphere.

Another important and interesting sink of  $NO_2$  is a reaction with peroxyacetyl derivatives. The simplest reagent in this class of molecules is acetaldehyde, which is produced both through vehicle emissions and naturally during the respiration of trees[87, p.320]. Acetaldehyde reacts with hydroxyl radicals and oxygen in the atmosphere to produce peroxyacetyl radicals:

$$CH_3CHO + OH \longrightarrow CH_3CO + H_2O$$
 (1.22)

$$CH_3CO + O_2 + M \longrightarrow CH_3C(O)OO + M$$
 (1.23)

The resulting radical combines with nitrogen dioxide to produce peroxyacetylnitrate, or PAN.

$$CH_3C(O)OO + NO_2 + M \longleftrightarrow CH_3C(O)OONO_2 + M$$
 (1.24)

PANs can refer to a large class of such aldehyde derivatives and all are formed in a similar way. The importance of this reaction lies in its energetics. At higher temperatures (ie. close to the ground), PAN exists in equilibrium with its constituents. However, when the temperature decreases PANs become much more stable. This mechanism can cause nitrogen dioxide to be transported over long distances, as a parcel of air containing PAN can move on a scale of several hours to nearby regions[88] before descending to the ground and releasing nitrogen dioxide. While the sources of pollution in cities have a far stronger effect on NOx concentration, PAN transport is important for introducing nitrogen dioxide into rural areas.

Transport of polluted air away from cities can be measured by comparing the combined concentration of ozone and nitrogen dioxide to the concentration of NOx[89]. A high abundance of ozone relative to nitrogen dioxide indicates that polluted air has arrived at a location [12, p.53] from elsewhere.

Thus, beyond the obvious importance of measuring the concentration of harmful gas species, broader information about the dynamics of the atmosphere can be deduced from instruments that can identify and measure the molecules involved in the Leighton relationship[89][12, p.53]. A distributed network of small sensors that could do this would be a major advance and allow transport models of pollution to be validated, as well as improving understanding of local and regional pollution sources.

## 1.5 Measuring nitrogen oxides

Environmental sensors can be divided into two broad categories: in-situ instruments, which sample from the environment immediately surrounding them, and remote sensing instruments, which can gather information at a distance either by passively collecting radiation, such as sunlight or the Earth's own infra-red thermal emissions, or actively sending out signals into the environment and measuring the response. The better an environmental sensor network can map pollutants in a region, the better scientists will be able through data analysis to attribute sinks and sources, predict potential exposure and inform city planning. Both active and passive sensors have disadvantages that must be taken into account to achieve this aim.

Remote sensors often suffer interference along the path between them and their target, which can be dealt with by characterizing (either though some other form of measurement or by modelling) the ambient conditions around the sensor and compensating for any known sources of error. Multiple instruments are required to map the locations of three dimensional signal sources.

In-situ monitors, by contrast, measure only the air directly in front of the sensor, thus how representative the sensor is of the surrounding environment can be a very significant problem. In some cases, such instruments also can produce incorrect information because the sensor itself interferes with the quantity that is being measured. A turbine anemometer is a good example of this, as the housing and support structure for the blades inevitably interferes with airflow. This kind of problem can be very difficult to overcome.

However, in-situ monitors do have relative advantages. For a start, remote sensors do not always give information in the same format as the variable you wish to measure; for example an open-path spectrometer gives column densities of gases but not actual molar concentration, and the latter is much more useful. Remote sensing, on the other hand, allows a single instrument to cover a wide area, and because fewer instruments are needed the individual units can be more expensive.

The concentration of  $NO_2$  can vary in urban settings on a time scale of seconds and a length scale of meters. For estimating exposure levels, any sensor must be able to integrate measurements of a rapidly varying quantity, and produce results with high enough resolution to get an impression of how pollutant concentrations change over time. Source attribution additionally requires measuring pollutants before they undergo chemical reactions[90]. To identify realistic avenues of research, the costs involved in deployment and operation of each measurement technology must also be considered, as this cost will determine the maximum number of sensors that can be employed for a given budget over a given time period, and thus the spatial resolution that a measurement campaign can achieve[91]. A campaign to characterize urban air quality may involve several kinds of instruments working simultaneously together.

#### 1.5.1**Optical spectrometry**



FIGURE 1.11: Diagram of a generic grating spectrometer. Collimated light enters aperture  $\mathbf{A}$  and is focused by mirror  $\mathbf{B}$  onto the diffraction grating C. The diffraction pattern is refocused by **D** onto an array of photodiodes (or a CCD) indicated by **E**. Signals from the detector  $\mathbf{F}$  must be interpreted by computer to determine from the diffraction pattern which frequencies are present.



A "spectrometer" is an instrument that determines the abundance of different components within a sample with respect to some variable. A mass spectrometer would take a material and determine the proportion of its constituent molecules, segregating them in terms of mass. A light-based spectrometer splits incoming light using a prism or diffraction grating, separating it into a spectrum of individual wavelengths, and then measuring the intensity of each part of the spectrum. Figures 1.11 and 1.12 show examples of optical spectrometers. Spectrometry is a diverse discipline, and visible light spectrometers of various types lie at the heart of many of the techniques for determining concentrations of gases in the atmosphere described in the following sections.

FIGURE 1.12: A simple optical spec-The prism splits light from a trometer. source, changing photon trajectories based on their wavelength - an incandescent lightbulb (middle) produces a different spectrum to a fluorescent bulb (bottom). Image credit [92].

## 1.5.2 Ground-based DOAS



FIGURE 1.13: Diagram of a DOAS instrument taking a measurement. Light from the sun A scatters as it passes through atmospheric gases B, and the light that is scattered along the line of the instrument aperture C enters the instrument D. The situation in reality is complicated by the curvature of the atmosphere, refraction and multiple scattering.

Differential Optical Absorption Spectroscopy (DOAS) is a technique used to determine concentration of gases that incoming visible and near-infrared light has passed through. DOAS instruments look for the characteristic absorption lines in the spectrum of a broad light source that has been shone through the target gas[93]. It can use either resonant cavities and internal light sources (closed path) or can have the spectrometer and light source in two separate locations (fixed path), measuring the density of trace gases between them. Fixed path DOAS instruments can use xenon arc lamps as a

source of light[94], but it's also possible to use ambient scattered light to make measurements (open path). Without the requirement to measure along a fixed line, open path DOAS can be a lot more flexible as a way of measuring gases over large areas.

Ground-based DOAS makes use of sunlight that has been scattered from molecules in the atmosphere, as shown in figure 1.13. DOAS instruments are capable of measuring multiple gases at the same time, but bad weather (heavy rain, snow and fog) can limit coverage[95, p.33]. While the first such measurements were made with instruments that looked directly upward (ie. zenith pointing), today MAX-DOAS (Multi-AXis DOAS) instruments are widely used. These open-path instruments either have a turntable that can change the elevation angle of a telescope that feeds into the spectrometer, or use a line-shaped aperture and specialised optics to measure multiple angles at once. The latter produce a two dimensional spectrographic image, with one dimension for the elevation angle, and the other different wavelengths within the spectrum.



FIGURE 1.14: (a) Cityscan atop a roof in London. (b) Example of data from a CityScan deployment in London[96], a heatmap of NO<sub>2</sub> concentration. The

"hotter" colour corresponds to higher concentrations.

The CityScan instrument is an example of a MAX-DOAS type instrument, with a turret fitted with a compact spectrometer (COMPAQS) which observes the atmosphere through a linear aperture[97]. A single CityScan instrument is capable of measuring column densities in a hemisphere extending from zenith to 5° below the horizon, out to a range of roughly 5-6 km, with column pixels that translate to a spatial resolution of 50m. Because of the combination of turret and aperture, when processed CityScan data can return a panorama of NOx concentrations (figure 1.14.

## 1.5.3 Airborne and satellite DOAS

DOAS techniques can also be used with light that has been scattered from the ground, from a spectrometer that looks down, rather than up as in MAX-DOAS - see figure 1.15. Ground level NOx concentration can be derived from these with methods of varying sophistication, for example making use of chemical modelling software (such as GEOSChem in the case of OMI[49]) or comparison with in-situ monitoring[98].



FIGURE 1.15: Open path DOAS using a downward-pointing spectrometer.
Light from the sun A is reflected off the ground B and into the spectrometer on the air or spacecraft C. The gas concentration E within the column D formed by the path of the light from the ground to the spectrometer can be retrieved.

This calculation is aided somewhat by the fact that higher concentrations of NOx in the atmosphere generally have anthropogenic sources close to the ground, and the lifetime of NOx in the atmosphere (hours) is much shorter than the typical mixing time between the troposphere and stratosphere (of order a few months)[99], and consequently a column with significant NOx will have most of this density contributed by NOx at ground level and in the boundary layer. Still, the retrievals can differ

significantly from ground-based in-situ sensors. In urban environments, OMI has a consistent seasonal bias, reporting concentrations around 17% lower than GEOSChem during summer and up to 36% lower during winter, and this is likely a consequence of limitations in the retrieval itself[49].

Several satellite missions have carried instruments capable of retrieving ground level NO<sub>2</sub>, among them:

- The Global Ozone Monitoring Instrument (GOME) operational from 1995[100], launched aboard the ERS-2 satellite.
- The SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY)[101] aboard ENVISAT which launched in 2002 and went dark in 2012.
- The Ozone Monitoring Instrument (OMI)[49], launched on the Aura spacecraft in 2004.
- GOME-2[102], carried on the METOP series of satellites since 2006.

Environmental monitoring satellites are usually placed in orbits that allow them to achieve full coverage of the Earth, but the time over which a full map of the world's pollution can be obtained varies depending on the mission and instrument - GOME achieved full coverage every 3 days, OMI is capable of mapping the Earth every day. Spatial resolution also tends to be very low for satellites. GOME pixels are 40 km by 320 km wide. OMI has a better resolution at 24 km by 13 km, but this is still large enough for a single pixel to completely encompass cities and prevent satellites from being able to attribute particular pollution sources.



FIGURE 1.16: Airborne air quality mapper, the precursor to ANDI, aboard an aircraft operated by BlueSky, with data from over Leicester[96].

Spectrometers mounted on aircraft can make open path DOAS measurements using backscattered light from the ground in much the same way as satellite instruments can. While aircraft are capable of mapping pollution at meaningful resolutions, a key limitation is that the flight is necessarily a snapshot of the level of pollution - planes must land eventually, making Airborne DOAS more a tool for confirming predictions and identifying hotspots than long term study.

The AVIRIS (Airborne Visual/Infra Red Imaging Spectrometer) series of spectrometers have been flown in various aircraft for the last 40 years[103], producing a very large dataset that can be used for retrievals of a variety of gases, and has recently had its specification updated to the "next generation" AVIRIS-NG. The instrument aperture cuts perpendicular to the flightpath. The Facility for Airborne Atmospheric Measurements (FAAM) aircraft can also be equipped with DOAS-capable spectrometers[104]. A Leicester-based example is the ANDI instrument, which makes use of the same spectrometer as CityScan and can retrieve  $NO_2$  with a resolution of 80 m by 20 m on the ground[76].

### 1.5.4 Chemiluminescence gas sensors



FIGURE 1.17: Diagram of the operation of a chemiluminescence sensor. Intake air is split at  $\mathbf{A}$ , one stream of which is put into a catalytic converter  $\mathbf{B}$  that reduces NO<sub>2</sub> to NO. The valves at  $\mathbf{C}$  control which stream enters the chamber  $\mathbf{E}$  and reacts with ozone produced with the UV lamp at  $\mathbf{D}$ . The photons emitted by this reaction are detected with the light sensor  $\mathbf{F}$ , and the gas is then exhausted through the pump  $\mathbf{G}$ .

The principle of this detection method is fairly simple: Some chemical reactions produce light, and if a reagent exists that does so for the target gas, and is introduced in excess to it, the luminescence produced by the reaction will be proportional to the target gas concentration. Implementing a sensor based around this is more complex. Luminol has been used as a reagent in nitrogen dioxide detectors for some decades [105], using a nebulizer and photodiode to minimize detector size and filtration to make the oxidising reaction more specific to NO[106]. In spite of being a rapid and sensitive means of detection, there are several factors that

must be corrected for which make it less reliable[107].

Most modern chemiluminescence monitors do not use luminol however, instead relying on the photons produced when NO reacts with  $O_3$ . In about 8% of reactions, this produces NO<sub>2</sub> in an excited state, which, if it isn't quenched through collisions with other molecules that relieve it of its energy, can return to the ground state by emitting a photon[108].

$$NO + O_3 \longrightarrow NO_2^* + O_2$$
 (1.25)

$$NO_2^* \longrightarrow NO_2 + hv$$
 (1.26)

Reaction 1.26 produces a broad selection of photon wavelengths, peaking at 600 nm and stretching into the infrared with a second peak at 1200 nm.



FIGURE 1.18: A typical modern commercial chemiluminescence monitor, from Ecotech. The intensity of the light is affected by pressure and temperature, but both of these can be controlled in a sensor, an example of which is shown in figure 1.17. These instruments alternate between detecting solely NO in a sample, and converting the NO<sub>2</sub> in the sample to NO to give a combined NOx reading, from which the NO<sub>2</sub> concentration can be derived by subtraction[109]. Several methods of conversion are used in commercial instruments, such as photolysis

or a heated molybdenum catalyst. These have varying levels of efficiency, and don't necessarily only convert NO<sub>2</sub>. Molybdenum catalysts suffer interference from HONO, in a fairly well characterized manner which can introduce a typical bias in these instruments of around 2%. Photolytic converters can produce peroxy radicals from the concentrated VOCs found in exhaust gas samples, which can reduce NO in the NOx channel and cause the instrument to infer negative NO<sub>2</sub> concentrations, and due to the unpredictable nature of these reactions they are harder to compensate for[110]. Cross sensitivity with different chemical species can occasionally cause much more severe inaccuracies in NO<sub>2</sub> concentration[111].

The European Union and the United States Environmental Protection Agency defines chemiluminescence sensors as the standard reference instrument for continuous air quality measurements of NOx and  $NO_2[112]$ .

## 1.5.5 Cavity Attenuated Phase Shift Sensors



FIGURE 1.19: Diagram of a CAPS instrument. The air sample is drawn through the particle and humidity filter  $\mathbf{A}$  to the cavity  $\mathbf{B}$ . A lamp  $\mathbf{C}$  shines through this cavity, bouncing between the mirrors  $\mathbf{D}$  before leaking through to the detector  $\mathbf{E}$ . The cavity is drained using the pump  $\mathbf{F}$ . Representation of the source signal (intensity vs. time) at  $\mathbf{C}$ and output signal at  $\mathbf{F}$  are also shown.

These sensors (CAPS) work through different principles to the other optical methods described above. While they also shine light with a known spectrum into a target gas, instead of measuring the extent to which the gas prevents particular frequencies from reaching a detector, CAPS detects the phase delay caused by the photons interacting with the gas[113], as in figure 1.19. The cavity increases the effective path length of the photons through the sample chamber, making the phase shift effect easier to detect.

This type of sensor has proven very effective at measuring NO<sub>2</sub> specifically[111], using a light source with a wavelength of 440 nm, with a narrow bandwidth of 20 nm. Such optical instruments are simpler than a full spectrometer, and don't have the same problems with chemical interference as chemiluminescence. Although some 1,2 dicarbonyl species can absorb on similar wavelengths, this can be compensated for with an appropriate optical bandpass filter. The solid-state optical systems are also relatively easy to maintain. Commercial instruments with similar form factors to rack-mounted chemiluminescence monitors are currently available[114][115].

## **1.5.6** Advanced spectrometry techniques

Two particular types of spectrometer are used make precise and sensitive measurements of gases in the context of air pollution.



FIGURE 1.20: Diagram of an FTIR gas spectrometer. The beam from light source  $\mathbf{A}$  is split by  $\mathbf{C}$ , and one of the resulting beams bounces from the fixed mirror  $\mathbf{B}$ . The other shines through the sample chamber  $\mathbf{D}$ , and bounces off the movable mirror  $\mathbf{E}$ . The beams merge again at the splitter and are detected by the sensor  $\mathbf{F}$ . The computer  $\mathbf{G}$  controls the mirror position, and uses a Fourier transform to reconstruct the spectrum of light from the position-dependent in-

terferogram.



FIGURE 1.21: Diagram of a BBCEAS cavity arrangement. The light source **A** shines through the mirror **B**, and is then repeatedly reflected through the sample chamber **C**. The mirrors are not perfectly reflected and photons from the source light can eventually escape to the spectrometer **D**.

Fourier transform infrared spectroscopy uses mid-infrared light sources, can be used to measure solid, liquid of gas composition, and can use either an open or closed path and sample chamber[116]. It incorporates a Michelson interferometer which can very precisely select a light frequency, and uses computer post-processing to calculate its attenuation by the sample - see figure 1.20. FTIR spectrometers are expensive and bulky, but have good resolution and accuracy, and more advanced models are capable of giving results at very rapid intervals[117].

Broad-Band Cavity-Enhanced Absorption Spectroscopy (BBCEAS) instruments use a "resonating cavity" as the sample chamber. With two highly reflective mirrors either side of the gas sample, the photon path length through it can be greatly increased, leading to more chances for absorption and thus improved sensitivity - see figure 1.21. The "Broad-Band" refers to the broad spectrum of the light source; sometimes the name of the instrument makes it
clear that it the light is incoherent, as opposed to the more monochromatic light from a laser[118], which cannot be easily tuned to different frequencies of interest. Such systems have been used for a wide variety of gas[119] and liquid[120] sensing applications besides atmospheric chemistry[121]. Most BBCEAS instruments use a diffraction grating spectrometer as in figure 1.11, but some make use of the Michelson interferometer of an FTIR, giving even greater resolution.



1.22:FIGURE А laboratory Michelsonequipped BBCEAS system built by University College Cork in Ireland's Laser spectroscopy  $\operatorname{group}[122].$ 

While they work on different optical principles, both BBCEAS and FTIR instruments are capable of high levels of precision, specificity with regard to target gas species identification and low detection limits[119]. They can take measurements very rapidly, making them ideal for examining short-lived atmospheric species[118]. Their ability to unambiguously determine concentrations of gases through the structure of their absorption lines make them a gold standard, in a lab setting at least. The downside is that FTIR and BBCEAS systems are expensive specialist equip-

ment that is often assembled by universities and companies themselves for a specific purpose, and though commercial systems are available they are large, complex instruments that require training to operate.

#### 1.5.7 Metal oxide semiconductor sensors



FIGURE 1.23: Diagram of a MOS sensor. The air outside the housing **A** filters through the aperture to the semiconductor element **B**, which is heated by **C**.

Modern semiconductor (MOS) sensors such as the SGX MICS series [123] are small and lightweight, and simple enough that they cost only a few pounds - see figure 1.23. They are well suited to air quality applications[124] and have been used in environmental monitoring to detect  $NO_2$ [125], VOCs [126, 127], ozone [128] and CO [129, 130]. Although semiconductor gas sensors are very precise (of order a few parts

per billion)[131][132], they have long had issues with calibration drift over time and cross-sensitivity to different gas species[133].

Still, metal oxide semiconductor sensors have been in use for years in industrial applications [134] and as hydrogen safety sensors [135][136]. In safety applications, low frequency shifts in baseline concentration can be ignored as artifacts of sensor drift, but a rapid increase indicates danger and consequently the broad responsiveness to different gases can be a blessing in this field. But it's a curse for air quality monitoring, where absolute figures, specific to an individual component like NO<sub>2</sub>, with characterized uncertainties are required.

For examples of instruments that use MOS sensors, see section 2.6 on page 74.

#### 1.5.8 Electrochemical Sensors



FIGURE 1.24: Diagram of an electrochemical gas sensor. The sample air A passes through the grille to the working chemical electrode B. Isolated from this air is a reference electrode C which is used to compensate for temperature changes and leakage currents. The electrodes feed into amplifier circuitry through wires D and some types are biased - that is, a voltage is applied across them - through E.

Electrochemical sensors work like a battery with one terminal connected to an amplifier circuit and the other terminal being the atmosphere, as in figure 1.24. An electrode oxidises or reduces the target gas, and the sensor's electronics detect the resulting current, which is typically on the scale of picoamps[137]. While small and relatively inexpensive, these sensors were initially developed for measuring gas concentrations ranging from 1-10,000 parts per million. Most atmospheric pollutants have their concentrations measured in parts per billion, and at these very low levels interfer-

ence from humidity, temperature and trace gases become very significant [138] [133].



FIGURE 1.25: A commercial electrochemical sensor from Alphasense.

Power is supplied to the sensors in the form of a very stable voltage that biases an electrode inside the device. The only power loss is due to leakage currents, and the power needed to keep the sensors operational is a few nanoamps, typically overshadowed by the requirements of the amplifier circuits[139]. The currents they produce are also very small, and designing support electronics that will not inject power supply noise or absorb electromagnetic interference that would overwhelm the signal is a technical challenge. The sensors require several hours to stabilize after voltage

is first applied to them.

For over a decade, these sensors have formed the core of many successful air quality monitoring instruments. See section 2.6 on page 74 for examples.

#### 1.5.9 Diffusion tubes



FIGURE 1.26: A NOx diffusion tube sold by Gradko International.

Developed in 1976, diffusion tubes are a simple and inexpensive way to take in-situ measurements of nitrogen dioxide levels[140]. Their simple design and low cost is an advantage and has led to widespread deployment across the world, but measurements consist of the average concentrations over the course of several weeks. The extremely long time between data points make diffusion tubes helpful for identifying potential urban hotspots and seasonal variation[141], but not so much for identifying transient pollution events[142]. Still,

surveys using diffusion tubes are ongoing even today, although their sensitivity to wind pressure is a complicating factor in an otherwise very simple and reliable measurement technique[143]. They are viewed by DEFRA as being capable of less than 20% uncertainty of measurement over the course of a year if they are deployed properly[112].

# 1.6 The role of small sensors

Name	Type	Concentration		Resolution		Costs to		Cita
		Accuracy	Range	Time	Spatial	Deploy	Maintain	Uite
Space DOAS	Remote	10-35%	?	Daily	$\sim 20 \text{ km}$	A lot	Low	[49]
Aircraft DOAS	Remote	5-30%	20 +	Once	$\sim 1^o$	$\sim \! 500 \mathrm{k}$	Medium	[96]
MAX-DOAS	Remote	10-35%	20 +	$6 \min$	${\sim}20~{\rm m}$	${\sim}1\mathrm{M}$	High	[97]
Chemilum.	In-Situ	15%	2 +	2  sec	N/A	$\sim 20 \mathrm{k}$	Medium	[144]
FTIR	In-Situ	5%	0.1 +	$1 \min$	N/A	$\sim \! 100 \mathrm{k}$	Medium	[145]
BBCEAS	In-Situ	5%	0.02 +	$1 \min$	N/A	$\sim \! 100 \mathrm{k}$	Medium	[118]
CAPS	In-Situ	>1%	2 +	$1  \mathrm{sec}$	N/A	$\sim 20 {\rm k}$	Low	[115]
Electrochemical	In-Situ	${\sim}25\%$	10 +	$1 \min$	N/A	$\sim \! 10 \mathrm{k}$	High	[137]
MOS	In-Situ	${\sim}25\%$	10 +	$1  \mathrm{sec}$	N/A	${\sim}1{\rm k}$	High	[123]
Diffusion tube	In-Situ	?	?	4 week	N/A	$\sim 50$	Low	[143]

TABLE 1.1: Summary of different NO<sub>2</sub> measurement techniques. Unmarked concentrations are in units of  $\mu g m^{-3}$ . All the in-situ instruments except for diffusion tubes are capable of sacrificing time resolution to improve accuracy and range.

Remote sensors generally do not have adequate temporal or spatial resolution to be useful on urban scales. An array of ground based MAX-DOAS instruments can give a good indication of pollution sources in cities, with considerable uncertainty in absolute concentration. Satellites generally have pixel sizes that dwarf individual settlements, and airborne instruments are limited to specific missions and do not typically capture variation in a region throughout the day. And in every case, these remote sensing systems must be checked against networks of sensors on the ground to ensure they are giving realistic results[146], and these ground-based networks are the only observational way to fill the gaps left by remote instruments. Advanced optical detection methods like BBCEAS, CAPS and FTIR can supply NOx concentrations with excellent accuracy and repeatability, but their cost is a limiting factor in their deployment. While industrialized nations like the UK have invested in networks of high quality chemiluminescence sensors in air monitoring stations, areas with tighter budgets in poorer countries do not have the same resources, and these are areas where urban pollution can be at its most severe [147].

A solution to this issue is the use of a single reference base station to support a deployment of smaller, cheaper sensors. In the detection of  $NO_2$ , both Metal Oxide and Electrochemical sensors have questions about long-term reliability and drift, as they have been validated only on smaller timescales[91][50]. Neither small sensing technology can be used as a long-term reference comparable to chemiluminescence or spectroscopic instruments, at least at our current level of understanding, and should only be considered to give reliable outputs when they have regular opportunities for comparison with a reference source. However, the manual work involved in either redeploying or recalibrating the sensors must be taken into account when considering overall cost of a pollution mapping experiment, and may be significant compared to the low sensor unit cost.

Thus an experimental campaign of small sensors must be designed with their limitations in mind in order to be successful. They should work well as part of a limited duration campaign to understand the dynamics of air pollution in a specific, local region for the purposes of urban planning and pollution management, either as stationary sensors or as mobile units that can travel back to a central hub to be re-calibrated. The duration of these campaigns is every bit as important for practical use of small sensors as the sensors accuracy.

# 1.7 Research Objectives

In the course of this work, the author developed a new sensor platform for Metal Oxide Semiconductor gas sensors, the details of which are described in chapter 2. The research performed with these sensors permitted the literature to be built on in the following ways:

- As we shall see in chapter 2, there is no universally accepted "best" calibration technique for MOS sensors (although work has been done elsewhere toward this goal[91] as shown in section 3.2.3 on page 101), and there are many variables associated with such a procedure that can affect the quality of the sensor's data output. The overarching work of much of this thesis is to clarify the effectiveness of different methods of calibration (sections 3.2 on page 95 through 3.4.1 on page 120, and 3.4.3 on page 124 through 3.7.1 on page 143) for instruments carrying MOS sensors.
- Any sensor system can be influenced by its environment in unexpected ways, and MOS sensor instruments are no different. The problem is perhaps more severe than for most instruments, given the known issues with MOS sensor cross-sensitivity. To clarify the work already done with these devices, numerous environmental sources of error are considered, including the effects of the day/night cycle on sensor performance, and the possibility of gases that might cause cross-sensitivity issues being carried to the calibration site from specific sources that can be filtered out through examining the wind direction. These issues are discussed in sections 4.4 on page 166 through 4.5.2 on page 176.
- Chapter 3 describes existing analysis techniques that have been used for MOS sensors, which once again, are not universally agreed upon. Linear regression fits require a simple equation for matching sensor output voltages to gas concentrations, which is not universally agreed on. Beyond this, more advanced data analysis techniques, useful for improving performance of MOS sensor instruments, are given in sections 3.7 on page 137, 3.4.2 on page 120 and 4 on page 146 through 4.1. Machine learning techniques have also occasionally been used in the literature, although a detailed investigation into the best method for using artificial intelligence to fit data to reality is beyond the scope of this work, an overview of one technique and results producing it is included in appendix C.

- The physical features of MOS sensors require special consideration in both instrument hardware and experimental design. The effects of manufacturing variations, the time a sensor takes to stabilise, and the difference actively drawing air with a fan past the sensors makes to their response are all investigated in chapter 5.
- That variations in the sensor manufacturing process can produce MOS sensors with different sensitivity characteristics even through the same design has been known for some time. A large-scale systematic examination of this issue is made possible by the unique technology of the instrument developed by the author as described in section 5.3 on page 190.
- There have been several deployments of mobile air quality instruments described in the literature, which are reviewed in the next chapter. Building on these experiments with "roving" devices, a case study is described in sections 6.2 on page 209 through 6.2.3 on page 222, leading to discussion of representivity and the regional transport of atmospheric pollutants.
- The low cost of small sensors enable many of them to be run in the same place, at the same time. A case study shows one way of using this ability to attempt to pin down the rate at which nitrogen dioxide concentration in the atmosphere changes, in sections 6 on page 195 through 6.1.5 on page 208).

# Chapter 2

# Review of low-cost gas sensors and instruments

In this chapter the design of the sensor platform will be discussed, beginning with a comparison of existing hardware components. There is a definitional distinction that will be important in this work:

- *Sensors*, referring to components that detect environmental variables and change their output accordingly.
- *Instruments*, which are platforms containing one or several sensors, any electronics or hardware features needed to support them, as well as a means of recording or transmitting data.

Drawing firm conclusions on the capabilities of any one type of sensor is made more complicated by the fact that their performance is naturally affected by the instruments that mount them and, perhaps critically, the data analysis techniques used to calibrate them.

A review of calibration techniques will be presented later in section 3.2 on page 95. Here, a rundown will be made first of the individual sensor elements themselves, and then of the most prominent currently available instruments. What represents the cutting edge will doubtless change over the course of the next few years, just as it has over the last few in a rapidly evolving field. The lessons that can be drawn from existing instruments are enduring, and these informed the design of the bespoke equipment that facilitated this work even when it began four years ago.

The objective is to produce a list of essential features for a sensor platform that will be suited to the requirements mentioned in section 1.7 on page 47.

# 2.1 Metrics of performance

There are several valid criteria for evaluating the performance of a particular combination of sensor, instrument and fitting technique. The simplest is the standard deviation of the residual (equation 2.1),

$$\sigma_{xr} = \sqrt{\sum_{i}^{N} (x_i - r_i)^2} \tag{2.1}$$

where x is a predicted value and r is a reference value.  $\sigma_{xr}$  gives a threshold difference between the sensor's prediction and the reference that the sensor is capable of meeting 68.2% of the time.  $2\sigma_{xr}$  this gives the range of values within which roughly 95% of the sensor's data points lie, and this latter metric is more rigorous for determining typical instrument performance. It is referred to in this work as RSE, residual standard error, although occasionally the undoubled standard deviation is referred to explicitly. RSE is a straightforward measure of the typical difference between predicted and actual concentrations, but for situations where the reference instrument measures extremes of concentration, residual standard error can be misleading - for example, an RSE of 50 µg m<sup>-3</sup> represents a more forgivable error if the atmosphere at the time contains 500 µg m<sup>-3</sup> of a target gas, as opposed to say  $10 \,\mu g \, m^{-3}$ . An alternative is fractional error (equation 2.2),

$$\delta_{FE} = \sqrt{\sum_{i}^{N} (\frac{x_i - r_i}{r_i})^2}$$
(2.2)

, where a fractional error of one or greater means the sensor predictions differ from the reference by an amount that is comparable to or larger than the actual concentration of gas in the atmosphere. Fractional error complements RSE as a measure of fit quality at a particular point. Another measure is the Pearson Correlation Coefficient (PCC) (equation 2.3),

$$\rho(x,r) = \frac{1}{N-1} \sum_{i}^{N} \frac{x_i - \overline{x}}{\sigma_x} \frac{r_i - \overline{r}}{\sigma_r}$$
(2.3)

where  $\overline{x}$  is the mean and  $\sigma_x$  the standard deviation of the predicted values. This value returns between zero and one for positively correlated values. The closer to one it is, the closer the reference and predicted gas concentrations are to moving in concert. A good *PCC* score means the instrument is sensitive - increasing its predicted concentration when the gas becomes more concentrated, and vice versa - but not necessarily accurate - higher *PCC* doesn't mean there is not a linear offset between prediction and validation, unlike the *RSE* and *FE* statistics.

For most of the scientific literature, sensor performance is given in terms of absolute error (measured in this thesis in micro grams per meter cubed) or correlation between a sensor's output and a reference. Since this is a univariate comparison, correlations given in the literature as coefficient of determination ( $R^2$ ) are converted to *PCC* by taking the square root. Occasionally other statistical metrics are used that do not directly convert into these standards; these will be noted when used, although direct comparisons are different.

An additional important metric is the detection limit of a sensor. This is the point above which a specific substance becomes distinguishable from no substance. By convention, the instrument detection limit is the level of substance that produces an average signal equal to three times the standard deviation of the "noise" output of the instrument, that is, the signal produced when none of the target analyte is present. The detection limit for an entire experimental method incorporates the instrument detection limit, along with any introduced noise due to dilution or other complications from sample preparation.

#### 2.1.1 Accuracy requirements

The regulatory demands on a sensing system are quite complex in Europe and vary from region to region. As mentioned in section 1.5.4 on page 38, the standard reference instrument for NO<sub>2</sub> and NO is a chemiluminescence monitor system, with typical specifications given for linearity (1-2%), lower detection limits (less than  $2 \,\mu g \, m^{-3}$ ), full-scale precision (1%), and 95% response time (less than 180 seconds)[112]. This may not be appropriate for small sensors, which as low cost devices can be designed for a different level of performance.

As mentioned in section 1.4.1 on page 29, NO<sub>2</sub> and O<sub>3</sub> concentrations change in a distinctive daily cycle. Figure 2.1 shows this in data taken from the AURN station next to the University of Leicester (number UKA00573), an urban background site. The peaks in NO<sub>2</sub> concentrations correspond to the increased traffic during rush hour, although this AURN station experiences generally lower NO<sub>2</sub> than most urban background sites[17]. Except in rare episodes, NO<sub>2</sub> in urban settings in Europe varies between 5-90 µg m<sup>-3</sup>[24]. The World Health Organization's European guidelines[149, p.179] for NO<sub>2</sub> suggest a maximum one-hour threshold of 200 µg m<sup>-3</sup>, justified by the effects acute exposure to the gas can have on asthmatics, and a yearly average of no greater than 40 µg m<sup>-3</sup>.

The data quality objectives of EUROAIRNET[150], a project by the European Environment Agency to produce a continent-wide sensor network, are expressed



FIGURE 2.1: Data from the University of Leicester AURN site showing concentrations of  $NO_2$  and  $O_3$  through 2016, per hour of the day. The central line is the median value, the shaded area represents the first and third quartile, and the dashed lines represent the 5th and 95th percentile. Maximum 95th percentile  $NO_2$  and  $O_3$  for this site in 2018 was  $52.0 \,\mu g \,m^{-3}$  and  $83.27 \,\mu g \,m^{-3}$  respectively[148].

in practical terms[150, p.25], derived from purposes toward which the data will be put. For example, if trend monitoring is desired, the required precision depends on the expected trend - of order 10% to 50% in either direction per decade. If the trend is to be detectable after a year, then an accuracy of 1-5% is therefore required.

The situation is different for small sensors, which should be seen as supplemental to more expensive reference instruments, rather than a replacement for them. In the case of these instruments, an RSE 95% confidence level of 20 µg m<sup>-3</sup> is enough to be able to discern representative levels of pollution - Since the upper regulation limit is ten times this margin, and the annual mean requirement is twice that, such a level of precision can be used to discern "good" days from "bad" with regard to NO<sub>2</sub> and O<sub>3</sub> concentrations that are hazardous to human health.

The time accuracy requirement for a mobile sensor is more exacting than for a

stationary device, where hour-long integration periods are the norm. A mobile sensor, while being used to map  $NO_2$  concentrations, will need to take several measurements per minute.

With regard to fractional error for small sensors, the ability to accurately report absolute NO<sub>2</sub> concentrations becomes important approaching the maximum onehour threshold rating of 200 µg m<sup>-3</sup> - but if the typical concentrations are as low as one tenth that level, then a lower fractional error threshold becomes acceptable. For the AURN urban background site at Leicester, the average level of NO<sub>2</sub> throughout 2016 was 27.5 µg m<sup>-3</sup>, and 95% of the measurements were within  $35.6 µg m^{-3}$  of this value. In the context of such low concentrations, a fractional error goal of 0.5 is required to make useful observations.

PCC standards have been already set, in a way, by the capabilities of the other instruments mentioned in this chapter. However, rather than having direct relevance to human health, PCC is a measure of how well an instrument tracks a reference - it refers more to the health of the instrument. With that in mind, a PCC comparable to the best-performing sensors is desirable, but not vital.

Finally, the timescale over which these standards must be achieved again depends on the application of the instruments. For mobile sensors, they might remain at the calibration site for almost all of the time before being deployed, but for small, stationary instruments, less frequent calibrations are practical. The length of time over which a particular level of measurement quality can be achieved is then a question of application. For the purpose of this work, results within one month and five months from calibration time will be considered.

#### 2.1.2 Uncertainty analysis

During the data analysis in this work, figures quoted to the right of a  $\pm$  sign are the 95% confidence interval of the quantity in question. Occasionally it may be

more appropriate to quote the error in terms of standard deviation, which is half the range of the 95% confidence interval. Unless stated otherwise it is assumed that the uncertainty in any result conforms to the normal distribution.

#### 2.2 Introduction to small low cost sensors

Recently the market for small air quality sensors has become quite crowded[151], corresponding with increased interest in the potential of such devices from governmental agencies for evaluating public exposure[24][151][152], as well as private citizens and NGOs[153]. Particularly targeted toward the latter are numerous crowd-funded projects[133] and startup companies that offer instruments that have varying levels of thought put into their quality and fidelity[154][155][156][157], as well as small sensors that can be integrated into homemade "maker" projects[158][159].

This proliferation is likely due to both a reduction in costs for relevant electrochemical and MOS sensor technology and easy to use, powerful microcontrollers like the Arduino[160][161], which coincided with the rise of the maker movement, crowdfunding and Hackspaces. These compete for attention with instruments made by more mature scientific companies and universities[91][50][162].

For all the diversity of approaches to instrument design, different levels of scientific expertise represented in the companies, and demographics targeted by marketing, a strong dividing line between instruments is the quality of scientific work supporting the manufacturer's claims of accuracy, sensitivity, and reproduceability.

What follows is a discussion of the sensor design for measuring  $NO_2$ , NOx and  $O_3$ . Only sensors and instruments for which there is available published information are used, and that are currently available at the time of writing - consequently this section should not be seen as a review of the current state of the art, but instead as the presentation of a representative assortment of currently used instruments, for the purpose of highlighting key design features and how they correlate with a particular sensor or instrument's level of performance. Performance statistics for sensors are translated, where possible, into the metrics defined in section 2.1 on page 51. The discussion first focuses on the kinds of small sensor elements available and the published evidence of their performance, moving on to the sensor instruments that are most frequently deployed and whose capabilities are best attested to in the literature. Mention will also be made of different approaches of commercial instruments that highlight the pitfalls of sensor design.

This discussion was informed by reports into the performance of small sensors produced by the EU[152], and the US-EPA[151], and a 2017 review of small sensor types[163]. These are cited specifically in the text whenever they are directly used, along with additional publications specific to individual sensors.

# 2.3 Review of available sensors

As discussed in section 1.5 on page 31, there are two types of  $NO_2$  and  $O_3$  sensor suitable for continuous measurement in small, low cost air quality monitors. They are distinguished by their mechanisms of action: electrochemical (ECM) and metal oxide semiconductor (MOS).

Numerous commercial sensors have been developed using these two technologies, and are marketed toward manufacturers of air quality instruments. The following section groups together related sensor elements with the same mechanism of action and manufacturer (but possibly different packages) are grouped together.

#### 2.3.1 Aeroqual SM50 O3



FIGURE 2.2: SM50 O3 daughterboard. The grey cylinder on the left is the sensor element.

The Aeroqual SM50-O3[164] sensor is available as a daughterboard for installation into larger instruments and can be fitted with several kinds of Metal Oxide Semiconductor sensing elements[165] (of which the OZU and OZL have a detection range appropriate for  $O_3$  air quality monitoring). It is also the basis of several instruments sold by Aeroqual[166] themselves. Such instruments have been deployed as part of urban air quality monitoring experiments for the last five years at least. The daughterboard features a digital output and an integrated fan, and is supplied cal-

ibrated from the factory.

As part of the Community Air Sensor Network (CAIRSENSE) project[167], several types of small sensors including the SM50 were tested in a small rainproof enclosure (roughly the size of a Leicester city council household wheelie bin) next to a regulatory monitoring site in suburban Atlanta, USA. For ozone measurements,

the SM50 performed extremely well, with *PCC* of between 0.95 and 0.98 compared to the Federal Equivalent Method instrument, a Thermo Fisher Scientific model 49i Ozone analyser.

Another study was performed in Haifa, a coastal city in Israel[168], mounted inside a CanarIT[169] instrument. The six sensors were calibrated next to an O342M analyzer (Environment S.A. LTD) with a precision of  $0.94 \,\mu g \,m^{-3}$  in a commercial district, before being installed in pairs at three locations near a busy road junction in city, roughly 100 meters apart. At one point during the experiment, one sensor from each pair was shuffled to a new location. Sensors from the same site exhibited high correlation with each other with *PCC* between 0.92 and 0.97. Unsurprisingly, sensors at different locations exhibited correlations as low as 0.4. However, even for sensors at the same site, the factory calibration returned concentrations for ozone that differed in absolute terms by up to  $40 \,\mu g \,m^{-3}$ , in spite of their good correlation.

The SM50, as the core of the Series 500[166] sensor manufactured by Aeroqual, achieved good correlation with an urban reference UV-absorption instrument in Edinburgh, UK[170] over two months, producing *PCC* values of 0.95 against a reference EnviroTechnology 400E photometer. It was found to have a small linear offset from the factory calibration that could be simply corrected after a deployment against a reference instrument.

The conclusion is that MOS  $O_3$  sensors can display a high degree of repeatability and are thus useful for determining spikes in the background levels of ozone, but factory calibrations cannot be relied upon alone, particularly for measuring absolute values of ozone concentration.

#### 2.3.2 SGX MICS series O3



FIGURE 2.3: MICS OZ-47 demonstration board

The SGX Sensortech MICS series of MOS sensors are available as standalone components without any support electronics, but a module with the required electronics is also available in the form of the OZ-47[171]. The individual ozone sensor element comes in several packages - MICS 2610, a legged component[172], MICS 2611, which has an identical sensor element to the 2610 except with a plastic protective cap, and the MICS 2614, which is even smaller than the 2610 and is designed to be surface-mounted[173]. Their extremely small size and low cost make them an attractive choice for Original Equipment Manufacturers (OEMs) of scientific instruments.

In spite of their forming the core of many instruments, after a recent purchase of the manufacturer by Teledyne, the MICS ozone sensors are no longer available[174].



FIGURE 2.4: MICS legged sensor

A study was carried out in the coastal city Aveiro, Portugal[175], with a suite of reference instruments installed in a van at a road intersection, side by side with commercial small sensors over a two week experimental period. The reference instrument for ozone was an Environnement O341M UV absorption instrument. Two instruments contained MICS sensors: the Libellium ISAG (no longer available), and the handheld AirSensorBox instrument (developed by the

EuNetAir group[176]). The AirSensorBox was only capable of a PCC of 0.36, and a root mean square error of  $16.1 \,\mu g \, m^{-3}$  for O<sub>3</sub>. In the same study, the ISAG instrument used the same sensors and scored a comparably disappointing PCC, and was out by some hundreds of parts per billion. By contrast, in Zurich, Switzerland, six custom sensor enclosures based around the OZ-47 module were deployed over three months at several urban air quality monitoring stations[177]. The reference stations use ultraviolet photometry and have an hourly uncertainty of  $4.1 \,\mu g \,m^{-3}$  at the limit  $120 \,\mu g \,m^{-3}$ [178, p.44]. The urban pollution in Zurich is moderate, as it has a service-based economy with little heavy industry and motorized traffic accounting for 47% of NOx emissions[178]. This sensor experienced a root mean square error of 6-10  $\mu g \,m^{-3}$  and *PCC* of 0.84 for O<sub>3</sub>, when calibrated using a linear equation that took temperature and humidity into account. The study conclusion notes that the sensors were accurate only in the first 1-3 months of operation, beyond which their response changed.



FIGURE 2.5:

sensor

sur-

mount

MICS

face

In 2014 a study in Denver, USA was carried out, using custom instruments fitted with MICS-2611 sensors[50], which were deployed at a central air quality monitoring station for two nonconsecutive weeks, separated by several months. The instrument, called the mPod, returned varied results compared to a reference station equipped with a Teledyne 400E ozone photometer, with residual standard error ranging from 8.2-29  $\mu$ g m<sup>-3</sup> for O<sub>3</sub>.

The large variation in performance between different MOS sensors underscores the importance of good instrument design in getting the best out of the fundamental components. At the same time, the best results from these studies hint at the potential of this class of sensor for reliable ambient air measurements.

#### 2.3.3 United Sens-IT



FIGURE 2.6: Sens-IT module, with inlet and outlet on the top and bottom.

The United Sens-IT[179] is a MOS  $O_3$  sensor module. It has an aluminium chassis containing a fan, support electronics for a calibrated digital output and the sensor itself. It is clearly intended, like the SM50, to be incorporated into a larger instrument (although it is much bulkier, being about the size and weight of a grapefruit[180]), and also arrives calibrated from the factory. It can be found at the core of the larger ETL-3000[179], a standalone air quality monitor marketed by the same company.

In spite of being available for several years, published studies on the Sens-IT module's capabilities are limited. The ETL-3000 has been tested in a closed gas cell[181], exposed to randomized concentrations of  $O_3$  of up to  $200 \,\mu g \,m^{-3}$ . It performed modestly, with a standard deviation at  $200 \,\mu g \,m^{-3}$  of  $6.6 \,\mu g \,m^{-3}$ . These results are not directly comparable to field tests, and all other measures of the Sens-IT's performance are produced by Unitec themselves.

#### 2.3.4 Alphasense OX-series



Alphasense produce many types of electrochemical sensor, and the OX series focuses on measuring ozone[182]. Unlike with MOS sensors, which can have a very small sensing element compared even to surface mount devices, the electrochemical sensors are a fuel cell within which a maction accurs with atmospheric respected produce a

FIGURE 2.7: Alphasense A-series

within which a reaction occurs with atmospheric gas to produce a very small current, which is proportional to the concentration of the gas in question. Because of this, the size of the cell affects its

performance characteristics[137], even between sensors that otherwise share the same chemistry. This series of sensors has been supplanted by the B431, but the

B421 sensors are compact, cost around a hundred pounds per element, and their modest power requirements and established reputation mean they are commonly used[152]. The most widely tested model is the OX-B421, although Alphasense only sells the similar B431 and A431 online. The B-series[183] has a greater surface area than the A-series[184] and is intended for use in larger instruments[137].

The OX-B421 was given a thorough investigation in a study published in 2017[185] as part of the AQMesh instrument. The sensor was subjected both to an artificial atmosphere in a gas cell with reference instruments, and field testing in Oslo, Norway next to four different air quality stations equipped with Teledyne API 400 UV photometry instruments for detecting  $O_3$ .



FIGURE 2.8: Alphasense

**B**-series

Two sensors were tested in the artificial atmosphere, and demonstrated a good linear response to the concentration of  $O_3$  present (*PCC* of 0.99 compared to the chamber's reference analyser), although the factory calibration was found to be off by up to 7.5 µg m<sup>-3</sup>, from an admittedly small sample size. The chamber concentration ranged up to 150 µg m<sup>-3</sup>.

The larger part of the study took place over six months. For the first three months, 24 AQMesh instruments were placed together next to a reference air quality monitoring station, before being distributed between three other stations in groups of at least four. Ten sensors remained at the first station to provide an assessment of their long-term performance.

The results of this experiment for the first three months are highly variable, with the best correlation between sensor and reference being 0.8 for  $O_3$ , but with eight of the ozone sensors giving a correlation of less than 0.5 (one as low as 0.09), and an average standard error of  $44 \,\mu g \,m^{-3}$ . By the authors' own admission, field calibration of such sensors remains a challenge.

#### 2.3.5 Citytech Sensoric O3-3E1F



FIGURE 2.9:

on a circuit board

O3-3E1F

sensor, mounted The Citytech Sensoric is a prominent, but less widespread series of electrochemical sensors, and the O3-3E1F[186] is available as an individual component for OEMs. In 2018 Sensoric have introduced an ozone sensor tailored for emissions testing[187]. The O3-3E1F ozone sensor has not been widely tested, but was part of a study in 2015 performed in Ispra, a small coastal town in Italy[188]. The study lasted six months, and involved placing a cluster of instruments close to the inlet of a reference air quality monitoring station

equipped with a Thermo Environment 49C UV photometer.

The purpose of the study was to evaluate different calibration methods, but in terms of comparative performance the Sensoric devices achieved a PCC of 0.92 to 0.97 between the predicted and reference datasets during a two week calibration period, and 0.76 to 0.9 in the remaining eighteen weeks of the experiment.

#### 2.3.6 SGX MICS series NO2



FIGURE 2.10: MICS surface mount package

Like the ozone sensors by this manufacturer, the same sensing element is sold in many different packages. Two of these have been well tested in the literature: the MICS 2710[189], which is a through-hole enclosure containing a single element which principally responds to NO<sub>2</sub>, and the MICS 4514[190], a dual sensor that houses a NO<sub>2</sub> sensitive element with the same properties, and another which responds to hydrocarbons.

These NO<sub>2</sub> sensors have failed to perform consistently under field conditions. During the Aveiro study [175], the ISAG instrument returned an almost completely incoherent PCC of 0.14 and was once more out by several hundred micrograms per meter cubed for  $NO_2$  compared to the Environnement AC31M chemiluminescence reference.

In the Colorado study[50], these sensors produced a far better match from their two weeks of calibration as part of the mPOD instrument against a Teledyne 200E, with a median root mean square error of  $16.0 \,\mu g \, m^{-3}$ . During the calibration period the mean atmospheric NO<sub>2</sub> concentration was close to  $30 \,\mu g \, m^{-3}$ 

The MICS 4514-equipped sensors in the Ispra study [188] had modest PCC scores during calibration of between 0.72 and 0.88, and for the rest of the experiment they achieved a PCC of only between 0.1 and 0.12. The reference instrument was a Thermo Environment 42C.

Importantly the Aveiro study [175] returned much higher PCC for the sensors during two weeks of calibration (on average 0.86), than during the subsequent roughly eighteen weeks of validation (between 0.24 and 0.25). Even then, this was with a linear regression model that included a corrective term for ozone concentration, without which the sensor performed far worse.

In spite of a mixed record of performance in the field, the sensors exhibited a good instrument-to-instrument correlation, of between 0.93 and 0.98[50], suggesting that they are providing indicative measures of air pollution, and would benefit from more advanced calibration techniques incorporating corrections for  $O_3$  interference.

#### 2.3.7 Alphasense NO2 series

There are several currently available Alphasense electrochemical  $NO_2$  sensors. These differ in their physical size (and hence somewhat in their performance), and in their electrode configuration. The A1 and B1 series have three terminals and are recommended for higher  $NO_2$  concentrations. Of these the smaller A1 found use in the CitiSense instrument[151]. The A43F and B43F have four electrodes and are better suited to the low  $NO_2$  concentrations found in environmental monitoring. The A1 and B43F sensors have not been tested against each other under identical conditions, but nonetheless their performance is demonstrated in several experiments.

The US-EPA performed testing of the A1 series in a gas cell as part of a CitiSense instrument[151]. The reference instrument was a Thermo Scientific 42C chemiluminescence monitor. In such controlled conditions, the sensor achieved a *PCC* of 0.98. Mead et. al.[91], in a seminal paper, did the same with A1 sensors and their own bespoke support electronics, and achieved *PCC* of better than 0.99. The gas cell used carefully control of mixing ratios for reference NO<sub>2</sub> and purified air as the reference.

In field measurements by the same Cambridge University group[91], two bespoke instruments (complete with NO and  $O_3$  sensors from the same manufacturer) were installed with a roadside AURN station for three days. The AURN station used a Thermo Environment 42C instrument for measuring NO<sub>2</sub>. Differing from other field installations mentioned above, in this case the sensors used an inlet pipe and were stored indoors, reducing interference from ambient temperature variations. During this time, provided corrections were made using the other sensors, the NO2-A1 achieved *PCC* of 0.89 and 0.97.

Field measurements were performed at Hrkingen in Switzerland, during which instruments equipped with B43F-NO2 and B4-NO sensors were located at reference stations for a month[191]. The reference instrument was a chemiluminescence monitor capable of a measurement uncertainty of  $4.1 \,\mu g \,m^{-3}$  for concentrations close to the limit value of  $80 \,\mu g \,m^{-3}$ . While the paper examined several advanced means of calibration which produced better results, using simple linear regression the sensors achieved a *PCC* of 0.77-0.87 against for the urban background and urban traffic reference sites respectively, and these numbers are most comparable to other sensors. By contrast, during the Ispra survey [188] two B4 series sensors were placed on a demonstration board in a custom enclosure. The results were less encouraging in these circumstances, capable of only *PCC* of 0.33 and 0.47 during the two week calibration period, although developing the calibration model in this experiment was harder for NO<sub>2</sub> than for O<sub>3</sub> because of a narrow range of ambient concentrations during the calibration period.

In the Aveiro study [175], the B4 sensor (as part of the AQMesh instrument) performed with distinction, achieving a *PCC* of 0.94, and root mean square error of  $1.9 \,\mu\text{g}\,\text{m}^{-3}$  compared to a reference Thermo Scientific 42C. The AirSensorBox[176] performed very differently, to say the least, with a *PCC* of 0.24, and a root mean square error of  $15.0 \,\mu\text{g}\,\text{m}^{-3}$  for NO<sub>2</sub>.

This particular study is instructive because with, several different instruments utilizing the same sensor under identical conditions, it produced data of drastically different quality. It shows just how critical variations in instrument hardware design and software can be. Another lesson is that calibration becomes markedly better, even in the outdoor atmosphere, when efforts are made to ensure that a packet of air entering the calibration setup is split and arrives at the reference instrument and the device under test at the same time.

#### 2.3.8 Citytech NO2-3E50

A sister to the O3-3E1F in the Sensoric series, the NO2-3E50 is an electrochemical sensor tailored to the measurement of nitrogen dioxide[192]. This sensor has been tested alongside Alphasense devices.

In the Ispra study [188], two 3E50s were mounted on an evaluation board for calibration. A simple circuit that converts the raw current data from the sensor element into voltages that can be read into a datalogger is not a full sensor system, and thus not directly comparable to the Alphasense devices mounted in an AQMesh. Nonetheless the sensor performance was good, provided interfering gases were taken into account, for the calibration period with PCC of 0.87 and 0.74. The validation period was not nearly as good, with PCC of 0.28 and 0.24.

For the Aveiro study[175], the sensor was mounted in the ECN Airbox. This design used a patented filtering and humidity sampling system for the NO<sub>2</sub> sensor. Software problems cut out roughly one third of the captured data, but during the co-location experiment the instrument's advanced design performed well, returning a *PCC* of 0.94 and a root mean standard error of only  $4.4 \,\mu g \,m^{-3}$ . This latter value is inferior to the one obtained by the AQMesh. The average annual concentrations of NO<sub>2</sub> in Aveiro are 25.1  $\mu g \,m^{-3}$ .

## 2.4 Summary: What sensors to choose?

Table 2.1 on the next page shows an overview of the capabilities of the above sensors and their recorded sensitivity. The sensor element is the core of any instrument, but the principal general conclusion from the above review is that the exact implementation of the sensor can make or break instrument performance.

Both electrochemical and MOS sensors have shown potential for air quality monitoring, but neither have a decisive advantage in terms of accuracy, and the physical characteristics of such devices involve a very different set of trade-offs for each type.

Building an instrument that can be equipped with either type of sensor will increase instrument size and complexity, but would also allow for maximum flexibility, or the possibility in future work of hybrid designs.

Among MOS sensors, the MICS series seems the logical choice. They are by far the smallest and cheapest, and although reviews in the literature are mixed, these sensors aren't appreciably inferior to other available MOS devices.

Primary gas	Brand	Sensor	Type	Form	Cost	Performance	
O <sub>3</sub>	Aeroqual	SM-O3	MOS	Board	Medium	Good	
	United	Sens-IT	MOS	Module	High	-	
	United	OZ-47	MOS	Board	Medium	Good	
		MICS 2610	ICS 2610 MOS PTH Low		Low	Poor	
	SGX	MICS 2611	MOS	PTH	Low	Poor	
		MICS $2614$	MOS	SMD	Low	-	
	$\alpha$ Sense	OX-B421	ECM	PTH	Medium	Poor-Good	
	Citytech	O3-3E1F	ECM	PTH	Medium	Good	
$\mathrm{NO}_2$	SCV	MICS 2710	MOS	PTH	Low	Poor	
	JUA	MICS $4514$	MOS	SMD	Low	Fair-Good	
	Sanco	NO2-A1	ECM	PTH	Medium	Good	
	αsense	NO2-B4	ECM	PTH	Medium	Poor-Good	
	Citytech	NO2-3E50	ECM	PTH	Medium	Fair-Good	

TABLE 2.1: Table summarising sensor models discussed in section 2.3 on page 58 above. The types are MOS: Metal Oxide Semiconductor, and ECM: Electrochemical. Form abbreviations are PTH: Pin Through Hole, a legged electronic component. These tend to be larger than SMD: Surface Mount Devices. Cost: Low means less than ten pounds per unit. Medium means less than a hundred pounds. High means less than a thousand pounds. Performance refers to performance during calibration (next to a reference) in the field. "Good" means achieving a PCC of greater than 0.9.

Between electrochemical sensors, the most well tested were the Citytech and Alphasense series. Both have exhibited inconsistent performance depending on the design of the instrument around them, but both have proven that they respond to an extent to pollutant gases. One possible factor that has made Alphasense devices more prevalent is the Analog Front-End (AFE) boards, which use standardised electronics built by Alphasense. The highly sensitive electronics required to measure variations in the sensor output current have been noted as a possible cause of unreliability[175]. The AFE can be equipped with up to four sensors together and provide a compact and convenient analog output voltage, using an amplifier design built by the sensor manufacturer.

As referred to in section 1.5 on page 31, which detailed the different means of measuring nitrogen oxides in the atmosphere, a critical factor in the deployment of small sensors is cost. Reducing the deployment costs of sensors directly translates

to an improvement in spatial resolution for experimental campaigns that can then afford to deploy multiple instruments with a limited budget.

For this reason, MOS sensors will receive the most thorough examination in this work. Both MOS and electrochemical sensors have performed inconsistently in the field, which gives a great opportunity for understanding. But MOS sensors are far smaller, cheaper and more compact, and thus for practical reasons they are more attractive.

# 2.5 How MOS sensors work

It has been known for over forty years that heated metal oxides change resistance in response to the composition of the atmosphere around them[193]. For example, sensors using a  $\text{SnO}_2$  substrate are sensitive to a large number of hydrocarbons, and they have been used in industrial safety equipment and chromatographs[194]. Other MOS materials have been investigated, including Ga<sub>2</sub>O<sub>3</sub> and WO<sub>3</sub>, which have different gas sensitivity profiles[195][196], and additives can also change the sensitivity profile of the devices dramatically[197], although as mentioned previously all such sensors exhibit significant cross-sensitivity.

The resistivity of silicon-based semiconductor can be changed by doping the material, adding impurities that inject electrons or holes into the lattice. Metal oxides work differently, with changes in the lattice itself determining the availability of charge carriers. Oxygen deficiency in  $\text{SnO}_2$  causes it to exhibit n-type semiconductivity, as the oxygen vacancies accompany the formation of free electrons[198]. NiO with deficient metal atoms becomes a p-type semiconductor, with the electronegative oxygen withholding electrons from the lattice and producing holes. Such materials used for gas sensors can be doped or plated to fine-tune the sensor's response to target gases[199]. MOS sensors are manufactured with different types of porous surface geometry, from granules to nanotubes and fibres. The structure of a typical MOS gas sensor is shown in figure 2.11, and is based on the model of MOS sensor behaviour given in Naisbitt et al. [200]. This equivalent circuit model defines three regions of the semiconductor - the surface which interacts with gases in the atmosphere, the bulk which is unaffected by it, and the particle boundary which lies inbetween these two regions. The particle boundary is situated at a distance from any material exposed to the atmosphere equal to the *Debye length*: the distance into the sensor that chemical electrostatic effects can propagate, which is related to the material's physical properties.

The sensor will have a baseline resistance that is related to the bulk and particle boundary resistance, but because of the random geometry of the granular sensor surface[201], this baseline resistance will vary between individual sensors, as will the gas responsiveness[200].

When the sensor material is warmed to between 100°C and 500°C oxygen molecules will adsorb onto the surface of the semiconductor material, possibly splitting into negative oxygen ions if the temperature is high enough. On an n-type semiconductor, the oxygen molecules become ionized and draw electrons from the surface region, and the lack of charge carriers will reduce the sensor's conductivity. For ptype semiconductors, the electronegative oxygen molecules inject additional holes into the surface layer, making it more conductive than the bulk[199].

While different gases in the atmosphere react chemically in different ways to the semiconductor material, for the purposes of this work "reducing gases" are molecules that bond to the sensor surface and donate electrons into the lattice, which will lower the resistance of an n-type material and do the opposite for p-type.  $CH_4$  is a typical reducing gas. Oxidising gases like NO<sub>2</sub> will have the opposite effect[196][193]. Sensors are designed to respond most strongly to one or more "target" gases. The working temperature that gives the best response will vary depending on the material of the element itself, any impurities and the particular target gas[197].

Any sensor material may participate in reactions with a variety of different gas species in addition to the target gas. In ideal conditions these reactions are reversible and have an equilibrium point determined by the temperature of the sensor - however, some reactions (which vary depending on the chemistry of the sensor's active surface) may be irreversible and will poison the sensor, making it less sensitive[132][196]. The natural consequence of this system of reactions is that although the chemical properties of the sensor material and the temperature of the element might favour reactions with a particular target gas, cross sensitivity with many different species is unavoidable.

Although deployment of multiple different sensors can compensate for the crosssensitivity issues in calibration, it can't eliminate it. MOS sensors can thus be used only in situations where any interfering species can either be measured by another means, or they must be calibrated regularly and used in locations where the background levels of possible interfering gases vary in concentration more slowly than the target gases.

Modern monolithic MOS sensors have a built-in resistive heater for bringing the element up to optimum temperature, but the actual surface temperature is a function of the power dissipated by the heater and the ambient conditions. The heater resistance reduces over time[203] and this variation has a corresponding effect on the device's operating temperature, and hence selectivity, as the sensor ages. Ambient humidity can also influence the conductivity. Water vapour adsorbed onto the sensing layer can react and act as a reducing gas, mitigating the sensor response to oxidizing gases[132].

In addition to these environmental factors, an initial warm up period is required for the sensors to achieve chemical equilibrium with the atmosphere. When cold, volatile substances such as water vapour deposit onto the sensor surface[197][128].



FIGURE 2.11: Diagram of the various types of interaction between atmospheric gases and a MOS sensor surface. In the leftmost region the sensor is unpowered (and exhibits the base resistance). The three other regions of the diagram describe different processes that actually occur simultaneously to varying degrees. The sensor's output is the resistance across the whole of the sensor material, which forms a resistor network with contributions from both the bulk and surface regions (although the non-sensitive surface will have similar properties to the bulk). This model of the sensor material also explains the wide variation in base resistance between individual sensors of the same type, as the random nature of the surface geometry means an equally random network of resistances. This diagram is a two dimensional representation of a three dimensional material; in an actual sensor, the sensitive region is spread into the surface with a distance dependent on the geometry of the sintering. Image first published by the author in Peterson et al. [202].

The warm up time is the period after which changes in sensor response can be attributed to changes in atmospheric composition, an equilibrium that can be determined to have been achieved when sensor response no longer changes with time in a fixed atmosphere. This must be determined pre-deployment for each sensor. A compromise between precision (longer warm up time) and time efficiency (shorter warm up time) must be made - to inform this, an investigation into the warm-up time is detailed in section 5.2 on page 183.

# 2.6 Review of instrument designs

Beyond choosing the right kind of sensor element, the design of the instrument containing it is critical. Features that directly support the sensor, such as the quality of the electronics interfacing with it or the use of fans to circulate air past it, directly affect data quality. The ability to run off solar panels or transmit data wirelessly can make an instrument more flexible, and thus easier to deploy. There is a vast gulf between lab-made prototypes and complete commercial instruments ready for deployment by anyone who buys one, and building a device more useful to the latter demographic will allow a reliable sensor to be widely applicable.

Several instruments have been the subject of field air quality tests, and as we shall see, many contain vital lessons on clever instrument design. But there are plenty of instruments that fill the opposite end of the quality spectrum. Even instruments that fail to perform well can still teach lessons about making informed design choices, but unfortunately some of these instruments are being marketed to this day, in many cases producing unreliable and unverified data. A comprehensive review of Kickstarter air quality instruments is impractical here. Instead, as with the discussion of sensor elements, a selection of notable examples will be mentioned in this section.

#### 2.6.1 AQMesh

The AQMesh[162] is one of the most widespread small air quality sensors. It has been sold for nearly a decade and been through many revisions during that time. The essential features of the instrument are a pyramidal metal enclosure containing a number of sensors, along with a GPRS (General Packet Radio Service) transceiver that broadcasts data over the mobile phone network[204].

The definitive version of this instrument is equipped with a large battery and a cluster of electrochemical sensors. Sensors are exposed directly to the environment on the underside of the housing, and because of their extremely low power requirements, the system can run for over a year on its internal batteries, requiring neither external electricity nor a solar panel. Other version of the sensor can be fitted with MOS sensors, sound detectors and particle counters, some of which require a fan to draw air into the instrument, and all these power-thirsty devices require an external power supply. The sensor's electronics design is semi-modular, with a single main board carrying the principal sensors and support electronics, and a daughterboard can be fitted to supplement these sensors.

AQMesh is one of the few commercial sensors to have rigorous procedures in place to teach end users how to take accurate measurements, specifically with a sensor afflicted with changing performance over time. Their calibration procedures are standardised, and they make use of proprietary algorithms to convert raw data into gas concentrations[167]. The unit cost is close to £10,000, which is quite high, but includes data processing and experimental support services.

The AQMesh has been used in various studies [175][185][91][138], and is commonly equipped with Alphasense sensors for air quality monitoring. The diversity of the models available makes it hard to draw definitive conclusions about sensor design from AQMesh performance, as the exact hardware being tested is not always mentioned, but it is a concrete example of the kind of features expected in a successful sensor.

It should be noted that a recent version of the AQMesh has demonstrated improved sensitivity in field testing, but these results have yet to be published[205].

#### 2.6.2 Cairpol CairClip



The CairClip is among the most compact complete sensor instruments. A schematic of the device's interior is shown in figure 2.12. The single electrochemical sensor sits behind a filter, and is accompanied by humidity and temperature sensors. It features an extremely small fan, to ensure fresh air reaches the sensor.

FIGURE 2.12: The interior structure of the CairClip instrument with an electrochemical sensor. Image credit [206].

The CairClip contains an electrochemical sensor of unknown make in an extremely compact package, A number of different internal sensors are available,

although for the purposes of this work the important sensors are an NO<sub>2</sub> sensor, and a combined NO<sub>2</sub>/O<sub>3</sub> sensor that returns a signal proportional to the combined concentration of the two gases. These sensors showed promising results during reference chamber testing[207] with injected NO<sub>2</sub> concentrations up to 288 µg m<sup>-3</sup>, recording a *PCC* of 0.99, but during deployment in the Ispra survey it was a different story.

A running theme during this review has been the issues of cross-sensitivity which must be accounted for. During the Ispra deployment calibration [188] against a Thermo Scientific 42C, the two NO<sub>2</sub> instruments alone only matched the reference with a *PCC* of 0.24 and 0.46. With external reference instruments supplying O<sub>3</sub> data to correct the output from the CairClip, using a multiple linear regression, the *PCC* improved to 0.76 and 0.86, however during the longer validation period *PCC* reached no better than 0.14 no matter the circumstances.

A survey near Houston, USA[208] underscored this result by including the combined sensor. Part of the experiment was performed directly by citizen scientists, but at two different calibration sites (one in Denver, USA and one at a local airport near Houston) one each of the combined sensor CairClips and a solo NO<sub>2</sub> Cair-Clip were placed next to each reference station (equipped with Federal Reference Method instruments) for 24 and 30 days. During this period the reference sites reported very low NO<sub>2</sub> concentrations of  $10 \,\mu g \,m^{-3}$  on average, and a combined NO<sub>2</sub> and O<sub>3</sub> average concentration of  $74.5 \,\mu g \,m^{-3}$  and  $144 \,\mu g \,m^{-3}$ . After a linear regression was used to fit the output signal to the references, the fractional error of the combined sensor was 1.02 and 1.16 for each site respectively with a *PCC* of 0.88, whereas the solo NO<sub>2</sub> sensor managed a fractional error of 3.87 with *PCC* of 0.54.

The sensor is well engineered for ruggedness and its small size is impressive, but the difference in performance between the two models illustrates the importance of cross-sensitivity for such instruments. If the solo NO<sub>2</sub> sensors were less vulnerable interference from changing levels of  $O_3$ , it would likely be an extremely useful device. On the other hand, the combined oxidant gas (NO<sub>2</sub> + O<sub>3</sub>) model can return very useful data for distinguishing between local and regional sources of pollution as discussed in section 1.4.1 on page 31. Provided there is also a reliable way to measure NOx too, this might be the most exciting use of this sensor.
#### 2.6.3 Aeroqual Series 500



Aeroqual is a modular handheld sensor system [166], costing \$500. The Series 500 has swappable "heads" that allow different gases to be measured, and one of these is fitted with the SM-50 MOS  $O_3$  sensor as reviewed in section 2.3.1 on page 58. It features an integrated GPS antenna for location stamping data, which thanks to its small size and eight hour battery mean it has found a use in mobile air quality experiments [209][210] [163]. It can also be directly connected with a computer via USB to collect longterm static data.

FIGURE 2.13: Aeroqual Series 500 handheld pollution sensor. Image credit Aeroqual [166].

In the Edinburgh study [170] the instrument was mounted roughly two meters away from the inlet to the urban background air quality monitoring station at St Leonard's. When looking at ozone alone, the instrument was capable of PCC=0.95 against an EnviroTechnology 400E photometric reference. The factory calibration for the Series 500 instrument had a slight linear offset for  $O_3$  that could be easily corrected for. The Series 500 can also fit an NO<sub>2</sub> sensing head, which performed far worse using factory settings. However, when a linear correction combining  $O_3$ readings from another Series 500 was used, the NO<sub>2</sub> instrument achieved a *PCC* of 0.94 against an EnviroTechnology 200E chemiluminescence reference.

## 2.6.4 Unitec SENS-IT and ETL-3000



FIGURE 2.14: Weatherproof enclosure of the Unitec ETL-3000 instrument, which makes use of the SENS-IT sensor. Image credit S.R.L. [179].

Unitec manufacture both the SENS-IT OEM sensor module and the ETL-3000[179]. The former is bulkier than most of the sensors aimed at OEM instrument manufacturers, and the ETL-3000, while more complete, is among the most expensive instruments in this category. It is also necessary to re-calibrate it after a period of six months[163].

The ETL-3000 is a very complete instrument, as befits its size (52cm high) and cost (\$2250)[179]. It has a wide array of peripheral options, including a GPRS modem and solar panel to allow it to operate independently in the field, but it is intended as a permanent installation and by default takes external power. It is clearly marketed as a rival to

bulky and expensive reference instruments, and while it is almost certainly easier to install than a chemiluminescence sensor and boasts lower running costs[179], its weak performance reported in reviews since its release in 2015 implies that if someone has the space and facilities for a larger, more expensive permanent instrument, they're better to go for a reference instrument that uses a technique directly recommended by the EU and US-EPA standards[112].

#### 2.6.5 mPod



FIGURE 2.15: The interior structure of the mPod. Image credit Piedrahita et al. [50].

The mPod is an example of an academically developed small sensor from the University of Boulder, Colorado USA. First in development from 2009[211], the mPod was designed from the outset to support scientific measurements and has already been deployed as part of measurement campaigns by the University of Boulder[212]. Equipped with metal oxide sensors, it can return results for both ozone and nitrogen dioxide. Due to its academic background, several conclusions relevant to the use of such sensors have been published[50], including the importance of calibration at a specific location.

Technology from this instrument was incorporated

into the open-source U-POD instrument[213], which has a bill of materials costing around \$1000.

#### 2.6.6 ELM

The ELM[214] is produced by Perkin-Elmer, a well-established company in the field of scientific instruments with a history going back to 1937. Formerly known as the Canar-IT until 2014, the ELM made use of an array of metal oxide sensors to give readings on pollutant gases. However, it was withdrawn from production in mid-2016 after failing to achieve its accuracy goals. Despite a superficial similarity to the AQMesh, the ELM's choice of sensor component meant that it could not run off a battery for an extended time[215], and the problems with using MOS sensors proved insurmountable.

# 2.6.7 #AirQualityEgg



FIGURE 2.16: Air quality egg exterior. Image credit Chantrell [216].

The #AirQualityEgg grew from an open source project among makers, but was eventually commercialized. It is one of the most widely deployed open-source air quality monitors, and has a history of disappointing those who expect too much of it. After a deployment in Louisville USA of some one hundred #AirQualityEggs in 2014, the devices reported such eccentric levels of carbon monoxide measurements that they were branded "dangerously misleading" [217] by the executive of the institute that had paid for their installation. Nitrogen dioxide readings were also highly inaccurate compared

to reference instruments. The initial version of the air quality egg had no fan and a small opening in its base, causing it to accumulate stale air, and the sensors were not calibrated before being deployed. A newer version of the #AirQualityEgg is now in production which includes a fan and is calibrated at the factory in a closed chamber[218].

# 2.7 Summary

The market for small air quality monitors was crowded even five years ago, and the sensors mentioned in this section are representative examples of a very large number of products. The performance of the sensors mentioned in the preceding sections is summarized in table 2.2 on the following page.

			Gas detection		Mechanical			
Name	Cost	Detects	Sensor	Paper?	Portable	Fan	Power	Comms
AQMesh	10,000	Nx Ox Cx PM	ECM	Yes	Outdoor	Some	Long battery	GPRS
CairClip	?	$NO_2 O_3$	ECM	Yes	Yes	Yes	Battery	USB
Series 500	$\sim 400$	Nx Ox Cx	MOS	Yes	Yes	Yes	Battery	USB
ETL-3000	$\sim 2,000$	Nx Ox Cx	MOS	Yes	Outdoor	Yes	Battery Solar	GPRS LAN
mPod	N/A	$NO_2 O_3$	MOS	Yes	Outdoor	Yes	Battery	?
ELM	1,000	Nx Ox Cx PM	MOS	No	Outdoor	Yes	Mains only	GPRS
#A Q Egg	230	CO NOx	MOS	No	Outdoor	Yes	Mains only	LAN

TABLE 2.2: Table of sensor characteristics. Abbreviations: MOS=Metal Oxide Semiconductor, ECM=Electrochemical, LAN=Local Area Network through an ethernet connection. The gas symbols Nx Ox and Cx stand for nitrogen compounds, oxygen compounds and carbon monoxide/VOCs respectively.

From this market research, the most important aspects of small sensor design can be identified:

- Low unit cost The cheaper the sensor is to deploy and operate, the more that can be bought and operated on a given budget. This allows for redundancy and greater coverage.
- Portability Small sensors are ideally designed to be portable, including ruggedization, low weight and the ability to run independently and without power supply infrastructure.
- Weatherproofing Sensors that can be mounted outside or in adverse weather improve deployment options considerably.
- Accessibility Sensors that are easy to use as consumer electronics have a large market potential already, although the citizen science movement has its own problems with lack of rigor that can throw doubts on a sensor's suitability.
- Communications Data collection usually needs to be accessible online, if not in the hands of the operator through their phone.

But further than this, there are some traits that only a few small sensors possess, which make them stand out from the crowd:

- Rigor Regrettably, sensors targeted at the lay population can be made profitably without ever being anything more than a platform for delivering advertisements and harvesting user data. Obtaining a degree of reproducibility and accuracy should be a non-negotiable goal for any scientific instrument. In the current market environment, even ensuring that minimal standards of rigor are upheld will make a sensor notable.
- Flexibility Sensors that can be customized or integrated into larger systems are rare, but desirable beyond their capabilities for air quality monitoring.
- Tracking Integrating a GPS antenna will allow a sensor to be moved around without needing to be tethered to a smart-phone and opens up possibilities for mobile experiments.
- Ventilation Representative measurements of gases in the atmosphere require a way for air to reach the sensor element. Whether sensors require active ventilation (through a fan or pump) depends on their application and the kind of sensors being used, and some successful instruments make use of less complicated passive air flow systems. This is a question examined in section 5.1 on page 179 of this thesis.
- Endurance This stems from portability requirements. Requiring mains power is not ideal; battery operation is better; the ability to function off solar panels or have a long lifespan is best.
- Open source Almost all modern sensors make use of open-source devices for the development. Aside from a moral obligation to give information freely to the community which should weigh in the decisions of all academics, there are practical benefits inherent in publishing the internal workings of a sensor, in

that it expands the potential developer base for producing additional sensors and peripherals.

Three years ago, there was no small sensor available that fitted all of the latter criteria, and yet all of these features were either beneficial to a small sensor project or essential to the task of monitoring air quality outright. As a consequence, the author was assigned by the Air Quality group at the University of Leicester to design an instrument that could meet these requirements.

# Chapter 3

# Novel calibration and data analysis techniques for a bespoke environmental sensor

This chapter describes in detail the hardware of the new instrument platform developed in part by the author over a period of four years, and the fundamentals of analysing data from experiments involving it. Section 3.1 on page 87 outlines the development process and gives the current specifications of the sensor relevant to the rest of this work. More detailed technical information can be found in appendix A. The design of the sensor evolved in response to newer experimental conditions - a brief overview of what guided that evolution is also given in this chapter, and justifications for the changes in instrument configuration are cited and described throughout the text.

For more information on the version history of the sensor, see section A.2 on page 234.

With the essential details of the instrument as a background, a summary of calibration techniques is given in section 3.2 on page 95. The development of rigorous calibration procedures for SOGS-MOS sensor systems, and a description of the first data analysis tools that supported them, is described in sections 3.3 on page 105 and 3.4 on page 116. Performance results and a novel way of displaying changing sensor performance over time is shown in section 3.4.3 on page 124.



FIGURE 3.1: Map of Leicester city, with north is directly upward. Districts are labelled in orange and divided by white borders. The two AURN stations are labelled in blue, with the University of Leicester station in the Victoria Park district.

Figure 3.1 is a map of the city of Leicester, showing the location of the two AURN stations within the city limits. St. Matthews is one of the poorest urban areas in the UK. Victoria Park, Riverside and Abbey Park are large areas of greenery. North Aylestone, the Space Park and the Canals region around Frog Island are the

more industrialised sections of the city. Leafy Stoneygate is a relatively well-off low density residential area. The West End, Spinney Hills and Belgrave contain largely low income terraced housing, and Braunstone and Beaumont have similarly impoverished 1930s-era detached housing. Clarendon, Dane Hills and Aylestone are quite diverse with both low and high income houses. Working class families and those students who cannot afford the en-suite corporate debt traps live sideby-side in the terraces of Highfields and Evington.

Most of the data examined in this chapter comes from the long-term experiment (LTE), a single sensor installation on the University of Leicester AURN station[148], a designated urban background station. The LTE ran from February 2016 until June 2017, giving over 15 months of data from a single sensor, although not all the data during this time could be used due to maintenance and software errors (see section 3.3.3 on page 112). There is also a designated roadside site in St. Matthews next to the A594 inner ring road[219]. This has no  $O_3$  measurement instrument and so is not used in this work. Figure 3.2 shows the statistics for the average concentrations of NO<sub>2</sub> during the year at both of these sites, giving context to the typical concentrations likely to be experienced in Leicester and giving the comparison with the accuracy thresholds in section 2.1.1 on page 53.

# 3.1 Introduction to SOGS and Zephyr

Beginning in summer 2013, the Small Open General purpose Sensor (SOGS) was initially conceived as a general purpose air quality monitor, as well as a specific test-bed for MOS gas sensors. Developed by the author as part of their PhD research, it evolved with the activities of the air quality group into a more general and flexible platform that could be carried on vehicles as well as deployed long-term at a static site. The Zephyr instrument, derived from SOGS, is the



FIGURE 3.2: Data from both the University of Leicester (urban background) and the A594 (roadside) AURN sites showing concentrations of  $NO_2$  through 2016, per month. The central line is the median value, the shaded area represents the first and third quartile, and the dashed lines represent the 5th and 95th percentile.

commercialization of this sensor system by Earthsense Systems Ltd., and a prototype Zephyr served as the instrument enclosure. Over the last three years work has gone into developing reliable calibration techniques, non-linear fitting of sensor voltages to retrieve  $NO_2$  and  $O_3$  concentrations, building up to experimental deployments and full field campaigns.

The instrument's main board can be incorporated as a peripheral to larger systems or as a datalogger suited to long-term deployments using a standardized electrical interface. As well as MOS sensors, the SOGS platform has been deployed with air opacity sensors for measuring particulate matter, more sophisticated optical particle counters and electrochemical gas sensors. Various prototypes of the Zephyr instrument, built around the SOGS platform, were the sources of data for this analysis and they all consisted of a weatherproof enclosure, peripheral sensors, fans and optional batteries.



#### 3.1.1 Instrument architecture

FIGURE 3.3: (a) A schematic of the main board of SOGS with functional regions labelled. (b) SOGS daughterboard example configuration. This diagram shows base board, an expansion board (through the digital Raspberry Pi header) and instrument boards (through the analog headers). Multiple boards can be stacked through the digital header, and up to four different analog instruments can be connected.

The SOGS base board is shown in figure 3.3(a). It is essentially a 32-channel 16-bit datalogger (*SENSORS*) with SD card storage (*STORE*), alongside a programmable battery charge regulator, solar panel maximum power point tracker with controllable power outputs (*POWER*), digital I/O (*COMMS*) and a header for fitting peripherals. So far a GPS/GSM transceiver has been developed to work with SOGS and is used for measurements from vehicles, but in theory it can work with any device that is designed to work with a Raspberry Pi computer, or be incorporated as a daughterboard itself into larger instruments.

Sensors using SOGS have a common architecture, shown in figure 3.3(b). The SOGS base board is connected to "instrument boards" through either its analog or digital headers. The MOS instrument board design used to capture the data described in this thesis is shown in figure 3.4, and carries four gas sensors targeted at detecting reducing gases, oxidizing gases (including NO<sub>2</sub>), ozone and ammonia. The base board is equipped with a combined relative humidity and temperature sensor. Details of these sensors are summarized in table 3.1 on the following page.

Target parameter	Part number	Method of detection	Sensitivity	Range	
			СО	1-1000 ppm	
Doducing	SCV Sensortech	Dodow	$NH_3$	1-500  ppm	
reducing	MICS 4514	reaction	$C_2H_5OH$	$10\text{-}500~\mathrm{ppm}$	
gases	MIC5-4514	reaction	$H_2$	1-1000 ppm	
			$CH_4$	>1000  ppm	
Oxidising	SGX Sensortech	Redox	$NO_2$	0.05-10 ppm	
gases	MICS-4514	reaction	$H_2$	1-1000  ppm	
Ozone	SGX Sensortech MICS-2614	Redox reaction	O <sub>3</sub>	0.01-10 ppm	
			NH <sub>3</sub>	1-500 ppm	
	CCV Concertach	Dadar	$C_2H_5OH$	y Range 1-1000 ppm 1-500 ppm 10-500 ppm 10-500 ppm 1-1000 ppm 0.05-10 ppm 0.01-10 ppm 1-500 ppm 10-500 ppm 10-500 ppm 1-1000 ppm >1000 ppm >1000 ppm 2000 ppm	
Ammonia	MICS 5014	ready	$H_2$	1-1000 ppm	
	MICS-5914	reaction	$C_3H_8$	>1000  ppm	
			$C_4H_{10}$	>1000  ppm	
Temperature,	GE Measurement	Capacitative	Temperature	$\pm$ 0.3 $^o\mathrm{C}$	
humidity	and control CC2D25	polymer	%RH	2%	

TABLE 3.1: Summary of different sensors used in the instrument, with information from the manufacturer's datasheets. The Sensitivity column for the gas sensors lists species that are meant to have a significant, detectable interaction with the sensor, beyond the target parameter.

The diverse sensor array is intended to help reduce the problem of cross-interference.  $NO_2$  may be principally detected using the oxidising gas sensor, but calibrating it in the real atmosphere might require contributions from the other sensors.



FIGURE 3.4: Image of the layout of sensors on a SOGS MOS instrument board. Visible are the (a) set of sensors; the second set of sensors (b) is on the reverse side. The MICS4514 contains two different sensor elements.

The doubling up of each type of sensor allows one set to work as a backup in case a manufacturing defect or sensor corruption over time renders one set unusable. The amplifier circuitry outputs a voltage directly proportional to the sensor resistance.

#### 3.1.2 Other instrument hardware

The essential hardware used during all the deployments described in this thesis consisted of a waterproof outdoor casing equipped with a SOGS, MOS instrument board and two fans. The case fans can exchange the air inside the case in less than a second and are positioned such that they minimize cooling of the heated active sensors. Using the fans to draw air over the sensors is essential to proper operation of the instrument (see section 5.1 on page 179). The instrument can be fitted with a battery and a GPS transceiver for location and time stamping, both essential for a mobile experiment platform. While capable of using a GPRS transceiver to send back live data, the version of SOGS used with the instrument had a flawed power supply that would cause brownouts and disturb the sensor data.

The casing is quite versatile and spacious enough to contain several types of sensors; in figure 3.5 a cutaway of one of the Zephyr prototypes is shown with an Alphasense AFE and bespoke MOS sensors installed. Although the sensor electronics design evolved, the signal processing chain was had an identical design across all prototype versions in this thesis.

SOGS has the appropriate electronics to be powered with an external DC supply, battery or solar panels. When equipped with a MOS sensor and transceiver, a 7200mAh internal battery lasts approximately 12 hours when taking measurements every minute.

The MOS sensor instrument boards fit onto SOGS via a ribbon cable and are mounted on a plinth suspended from the sensor's casing wall. Throughout this thesis,



FIGURE 3.5: Cutaway of the prototype Zephyr enclosure, showing layout of vital components and interfaces.

the oxidizing gas sensor element of the MICS 4514 is referred to as the "OX" sensor, and the ozone sensing element in the MICS 2614 is referred to as the "O3" sensor. This notation provides a distinction between  $NO_2$  and  $O_3$  (the gases) and OX and O3 (the sensors).

#### 3.1.3 Amplifier development



FIGURE 3.6: Diagrams of SOGS amplifier circuits.

The sensor manufacturers (SGX[190]) recommended the circuit shown in figure 3.6(a), which has the response equation 3.1.

$$V_{out} = V_{5V} \frac{R_f}{R_s + R_f} \tag{3.1}$$

where  $R_f$  is the sensor balance resistor,  $V_{out}$  is the output voltage of the amplifier,  $V_{5V}$  is the input voltage (five volts in the SOGS instrument), and  $R_s$  is the resistance of the sensor element. The circuit is a simple buffered resistive divider, and its voltage output is not linearly tied to sensor resistance. By design it will use only a small section of the potential range of output voltage if  $R_s$  is much greater than  $R_f$ . From an engineering standpoint, it also has the weakness of passing a current through the sensing element that will vary depending on the sensor's resistance. Power dissipated by current through a resistor scales linearly with the resistance, so if  $R_g$  is much greater than  $R_s$  this will affect sensor temperature. Although the heating element dissipated more power by a factor of a thousand (of order a few dozen milliwatts), the sensor's response to different gas species varies with temperature in an unpredictable way and this circuit design could thus complicate matters significantly.

The circuit used in the initial design of the sensor (Version 0.1) mitigated both these issues, as shown in figure 3.6(b). The output equation for this amplifier setup is

$$V_{out} = V_{5V} \left( 1 + \frac{R_{g1}}{R_{g2}} \right) \frac{R_s}{R_f} \left( \frac{\frac{R_f}{R_s} - \frac{R_{b1}}{R_{b2}}}{1 + \frac{R_{b1}}{R_{b2}}} \right)$$
(3.2)

where  $R_s$  and  $R_f$  are the resistance of the sensor itself and the feedback resistor respectively,  $R_{b1}$  and  $R_{b2}$  are the "boost" resistors that add a specific voltage to the sensor output, and  $R_{g1}$  and  $R_{g2}$  are the "gain" resistors that cause their amplifier to multiply the sensor output voltage by a factor  $(1 + R_{g1}/R_{g2})$ . This equation simplifies to

$$V_{out} = V_{5V} \left( k_{gain} \right) \left( k_{shift} R_s + k_{boost} \right) \tag{3.3}$$

where k terms are constants depending on the resistor values -  $k_{gain}$  comes from the gain amplifier, and  $k_{boost}$  and  $k_{shift}$  are more complex. In addition to producing a linear output, this circuit pulls only the input bias current through the sensor element, which using modern CMOS circuitry can be on a scale of nanoamps.

The voltage output is the quantity recorded by the sensor's microcontroller, but the sensor resistance is the quantity that is physically linked to gas concentration. The inverse of equation 3.3 is equation 3.4 - Notably, the summing amplifier produces a voltage that is negatively related to the feedback resistor.

$$R_{s} = R_{f} \frac{R_{b1}}{R_{b2}} \left( 1 - \frac{V_{out}}{V_{5V}} \frac{1 + \frac{R_{b2}}{R_{b1}}}{1 + \frac{R_{g1}}{R_{g2}}} \right)$$
(3.4)

The actual range in resistance of the MOS sensors was unknown before the project began; on the initial prototype the resistor values were unset and chosen after the sensors had been acquired and measured with an ohmmeter. However, on the next version (0.2.3) resistor  $R_{b1}$  was made a 100 k $\Omega$  potentiometer balancing a  $R_{b2}$  that had a value of 10 k $\Omega$ , and the divider  $R_{g1}$  and  $R_{g2}$  was likewise replaced, providing a gain of up to 20. Maximising the gain was particularly important on version 0.1, as the microprocessor's onboard analog-digital converter operated at only 10 bits.

This amplifier system performed well. For the deployed sensors the resistance values were chosen after laboratory testing, and probably constrained the full possible range of the sensors operation in atmospheres whose composition differs significantly from that of Leicester, as drastically different sensor resistance could cause the amplifiers to saturate.

#### 3.1.4 Timeline of deployments

The data analysis, hardware development and deployments were worked on simultaneously, but in this thesis the work is presented in logical order, rather than chronologically. The dates over which experiments were performed, and the effects they had on the sensor design are summarized below, and links to further exploration within this work are given.

- January 2014: First attempts at deployment on the University AURN station using sensor version 0.1.
- March 2014: First calibration results using inverse voltage equation.
- July 2014 Designing phase of SOGS 0.2.6
- Winter 2014 A. Brundle's work with active/passive aspiration (section 5.1 on page 179)

- March 2015 Put orders in for SOGS version 0.2.9
- December 2015 A. Aujla's roving sensor experiments (section 6.2 on page 209)
- February 2016 Installation of the long-term experiment (LTE)
- September 2016 Multiple adjacent sensors deployment on the AURN station (section 6.1 on page 196)
- June 2017 Conclusion of long-term experiment.

The specific design details for the sensors used in this work are important context for their deployment history. See appendix A for more information.

# **3.2** Review of calibration techniques

The objective of a large portion of this thesis is to investigate as many practical questions on the use of MOS sensors for pollution detection as possible. This was a three step process: First, the basic response of the sensor equipment to changes in gas concentration was examined. Next, a calibration procedure was developed for the sensors in a single reference location, with as many complicating factors controlled for as possible. Finally field deployments were made to verify these calibrations work in practice for real ambient air measurements, particularly in the light of potential interference from different background gas environments.

The first step was to determine an approximate equation relating sensor response to a target gas concentration. When the project started, aside from some terse information about metal oxide sensor response provided by the manufacturer datasheets [173][190], there was no clear consensus in the literature as to what this response equation should be. An experiment was performed in Leicester[220], using a closed gas cell with a reference Broad-Band Cavity Enhanced Spectrometer (BBCEAS)[221]. This gave a provisional relation between NO<sub>2</sub> concentration and sensor output, and the results of this are discussed further in section 3.2.2 on page 99.

The analytical techniques presented in the following sections have a well-defined process. With data from the sensor and reference over a limited calibration period in-hand, the constants in the candidate equation are adjusted, like a pick adjusting the tumblers in a lock, until that equation's prediction of the concentration comes as close as possible to the reference - with the square of the difference between reference and prediction being minimized. This is "least-squares regression", also known throughout the rest of this work as "fitting".

The method of calibration is an important question, as is whether to use a controlled environment, or whether to deploy the sensors in the vicinity of a reference instrument exposed to the same ambient environment ("co-location"). Section 3.2.3 on page 101 discusses this issue and provides the rationale for the latter calibration setup at the University of Leicester AURN station.

Calibration essentially means comparing an instrument's output to that of a reference, but beyond that the experimental setup can very enormously. Various calibration setups for small air quality sensors were used over the last four years, with the procedure being rigorously defined from mid-2015 onward. The basic parameters for a setup are the exact arrangement of the sensors (Section 3.3.1 on page 107) and the duration of the calibration (see Section 3.4.2 on page 120). It's reasonable to examine also the averaging techniques used for calibrations, and the consequences of timestamps being offset due to clock synchronization issues (Section 4.5 on page 174).

Running a simple least-squares fit to make raw data conform to a reference is at the heart of the data analysis work presented, but there are opportunities to improve precision markedly with more complex analytical techniques. These are discussed further in chapter 4 on page 146.

#### 3.2.1 Background of linear calibration equations

A simple model for metal oxide sensors is to consider them as devices with a limited active surface area A. Gases which bond to this surface (covering area a) influence the resistance of the sensor  $R_s$ , through the function f which is different for each gas species.

The equation describing this model is:

$$R_{s} = R_{b} + R_{g} \sum_{i \in G} f_{i}(a_{i}([G_{i}]), T_{a}, T_{h}) + R_{err}$$
(3.5)

where  $G_i$  is a specific atmospheric gas and  $[G_i]$  its concentration, with  $a_i$  being a function giving the fractional quantity of the target gas currently adhered to active sites on the semiconductor surface, relative to the entire semiconductor surface.  $T_a$  is atmospheric temperature,  $T_h$  is sensor element temperature, the  $R_b$  and  $R_g$  terms are base and gas sensitivity resistances to be found, and  $R_{err}$ constitutes any error due to the sensor aging or becoming poisoned. f is a nonlinear function unique to every gas in the atmosphere, and determining its form is the objective of this section, and different types of metal oxide sensor will have different responsiveness to different gas species.

The *a* term is also unknowable through measuring resistance alone, but physically important, as each sensor element has a limited active surface area, and the equilibrium equations represent dynamic chemical processes [193]. Ideally when the sensor is hot, gases cycle fairly frequently between the atmosphere and the surface of the sensor, and this keeps the *a* term approximately proportional to the concentration of each gas. This approximation holds only when the characteristic timescale of every reaction is less than the timescale over which the atmospheric conditions next to the sensor vary. In reality each equilibrium reaction will have its own rate equation, and some gases may adhere permanently to the sensor surface, which will increase the *a* term for that gas (as more of the limited surface area is

taken up with that gas). The sum of all a terms is proportional to, and must be less than, the total  $A_{sensor}$ :

$$\sum_{i \in G} a_i < A_{sensor} \tag{3.6}$$

which means that any gases that latch on to the surface permanently reduce the device's sensitivity to anything else, and it is possible under high enough concentrations of atmospheric gases for the device to become saturated.

Clues as to the form of the f function for the two sensors responsible for measuring atmospheric O<sub>3</sub> and NO<sub>2</sub> come first from the manufacturer's datasheets for oxidising and ozone gas sensors[173][190], as shown in figure 3.7.



FIGURE 3.7: (a) MICS-4514 Oxidising gas sensor element nonlinear response to nitrogen dioxide.[190] (b) MICS-4514 Reducing gas sensor element response to various gases.  $R_s$  is the sensor's actual resistance, compared to the base resistance  $R_0$ .[190]

These plots suggest the following relations over the concentrations shown:

$$f_{NO_2} = k_1 [NO_2] + k_2 \tag{3.7}$$

and

$$f_{NO_2} = k_1 [O_3]^2 + k_2 [O_3] + k_3 \tag{3.8}$$

where the k terms are constants to be found experimentally. The constants will be different for equations 3.7 and 3.8.

#### 3.2.2 Cell Calibration

One of the earliest experiments, carried out as described in Grant [220] in January 2014, was used a closed gas cell to determine the sensor's NO<sub>2</sub> response, with all other factors controlled for. An airtight enclosure was lined with PTFE to minimize reactions between NO<sub>2</sub> and the walls of the chamber, and the instruments were set up within, so that they co-sampled the gas pumped through. NO<sub>2</sub> levels were controlled by injecting samples with known concentration, and varied using a permeation tube. A reference Broad-band Cavity Enhanced Absorption Spectrometer instrument (BBCEAS) verified representative NO<sub>2</sub> concentrations from inside the chamber. Experimental data was taken by a version 0.1 SOGS/MOS prototype (see section A.2.1 on page 234). The results are shown in figure 3.8(a).

While only NO<sub>2</sub> was involved in the experiment, this was nonetheless the first part of the nontrivial task of producing useful gas equations. Somewhat surprisingly, the response of the sensor to varying NO<sub>2</sub> takes a different form to that suggested in the datasheet (see equation 3.7). The inverse of the MOS sensor output voltage was proportional to the BBCEAS readings with a *PCC* of 0.87 over a range of  $500 \,\mu g \,m^{-3}$ . Notably the relation held even below the manufacturer's stated detection limit of  $100 \,\mu g \,m^{-3}$ . The empirical form of f derived from this experiment is:

$$f_{NO2}(V_{OX}) = \frac{k_1}{V_{OX}}$$
(3.9)

where the k terms are individual constants to be found for each equation. However, at low concentrations, systematic deviation from this simple formula can clearly be identified. At lower concentrations where the fractional error was higher, the



FIGURE 3.8: Results from the BBCEAS calibration experiment, showing both raw output voltage and inverse voltage of the same measurement. The experiment demonstrated sensitivity up to  $400 \,\mu g \,m^{-3}$  as shown in (a), although at high concentrations the voltages from the sensors becomes quite low. A black dashed line in a) divides the data into reference concentrations below or above  $100 \,\mu g \,m^{-3}$ , solid lines are the least squares fit from the sensor voltage to the reference concentrations for the low concentrations, and the inverse voltage to the reference for the high concentrations. Typical levels of, NO<sub>2</sub> in urban environments are not often higher than  $100 \,\mu g \,m^{-3}[149]$ , although peaks can be ten times as high. For concentrations less than this, both normal and inverse voltage can be seen as approximately linear. Graph (b) shows the data for inverse voltage at low concentrations.

data could be fitted with a linear equation:

$$f_{NO2}(V_{OX}) = k_2 V_{OX} (3.10)$$

This equation achieved a PCC of 0.73 for concentrations below  $100 \,\mu g \,m^{-3}$ . In the United Kingdom, nitrogen dioxide concentrations generally stay below this level, so for most cases the sensor will respond approximately linearly. At higher concentrations, predictions using equation 3.10 will start to underestimate actual NO<sub>2</sub> concentration.

The results from this experiment have some important caveats. While the instrument responded strongly to changes in humidity [220], this response was not thoroughly examined during this experiment. Furthermore a calibration of air quality sensors taking place under conditions where the sole control variable is the target gas can demonstrate sensitivity, but the real atmosphere is a very different environment with many potential sources of interference not represented in closed-cell experiments.

In spite of its limitations, calibration in a controlled environment is attractive if it can be carried properly out with a reliable source of zero-NO<sub>2</sub> air, because it can clearly illustrate the sensor's non-linear response to the target gas while eliminating variation in interfering background gases as a source of error.

#### 3.2.3 Calibration strategy: controlled cells or co-location?

The atmosphere is comprised of many different gas species at different concentrations. Taking a single gas i from equation 3.5, if the response function  $f_i$  is approximately linear, irrespective of temperature, and the area taken up by all gases on the sensor surface a is much smaller than A, the function can be expressed as:

$$f_i \approx k_i [G_i] \tag{3.11}$$

, roughly the product of the concentration of gas  $[G_i]$  and the responsiveness of the sensor to that gas, condensed into the constant  $k_i$ . Equation 3.5 then becomes:

$$R_{s} = R_{b} + R_{g} \sum_{i \in G} k_{i}[G_{i}] + R_{err}$$
(3.12)

The  $k_i$  terms, however, can vary by orders of magnitude for different gases[193][197]. The simplest way of viewing a sensor calibration is as the process of varying the concentration of a specific target gas in a controlled way, and measuring the change in the sensor's output in response. The size of that response  $\delta R_i$  is equal to

$$\delta R_i = R_g k_i \delta[G_i] \tag{3.13}$$

where  $\delta[G_i]$  is the change in target gas concentration. It can be deduced from equation 3.13 that for a gas to be detectable at ambient atmospheric concentrations, the product of the its concentration and the sensor's sensitivity to it kmust be greater than that of other gases, lest the desired signal be overwhelmed. This leads to problems with calibrating these sensors (and indeed any instrument that produces a single output variable) that are distinct from spectrometer-based instruments.

A single measurement in time with instruments like absorption spectrometers consist of a set of values that each correspond to the intensity of light within a particular narrow range of wavelengths. Gas concentration is derived by fitting its absorption spectrum onto this output. Since the number of wavelengths a spectrometer measures can be in the thousands and a target gas's spectrum only needs to be the most prominent at a subset of these, a spectrometer is capable of accurately resolving gases at low concentrations - indeed, spectrometric methods can give results for several gas species simultaneously, via the structured absorption of the different gases at different wavelengths.

MOS sensors, on the other hand, output a single voltage that is representative of the air in front of it, air which is made up of different gas species. When all this diversity is condensed down into a single variable, and that variable is studied for the response to a single gas, any response from the sensor could have an unknown number of attributions. In a closed gas cell, it is possible in theory to control for all gas concentrations, using inert zero air and varying only the amount of the target gas present. But this can give a false sense of security when sensors are deployed in the field. In practice, in a realistic atmosphere, the detection limits for a target gas with a single-output sensor like MOS are set not just by the sensor's hardware and support electronics but by the chemical composition of the rest of the atmospheric environment. This is a known issue with low cost gas sensors[222], for example electrochemical gas sensors have a detection floor for  $NO_2$  that is practically determined in part by the level of  $CO_2$  in the atmosphere, shown in figure 3.9.



FIGURE 3.9: This diagram of the proportion of a electrochemical sensor response that can be ascribed to NO<sub>2</sub> or CO<sub>2</sub> shows a practical detection limit of around 20-30  $\mu$ g m<sup>-3</sup>, depending on CO<sub>2</sub> concentration. The conventional detection limit is the point at which 99.6% of the response is due to NO<sub>2</sub>, which on the diagram is the line where the green colour starts to change to yellow. Atmospheric NO<sub>2</sub> tends to vary between 10-100  $\mu$ g m<sup>-3</sup>, so the interference from CO<sub>2</sub> is significant for this sensor type[222].

All this nuance is completely lost, however, when testing in a controlled atmosphere. Hints as to the correct calibration equation for these small sensors can be gained from such experiments, but practical calibration must take place in the same kind of environment as they will be deployed in, with realistic concentrations of the target gas competing with naturally varying background gases. This conclusion has been borne out in work done by other groups, particularly a study in Denver, USA[50], which found lower standard errors during outdoor calibrations next to a reference source, and concluded that lab calibrations added bias and made it difficult to reproduce the range of atmospheric conditions that sensors would be exposed to during real experiments.

In Cambridge UK, a series of experiments were done in 2013[91] with electrochemical sensors, which share the characteristics of cross sensitivity with MOS sensors particularly the NO<sub>2</sub> sensor, which during a calibration against an AURN station over five days exhibited significant cross sensitivity with ambient O<sub>3</sub> concentrations compared to a reference Thermo-Environmental 42C analyser. This interference can be compensated for to achieve a *PCC* of 0.94 to 0.95, but only if atmospheric concentrations of those interfering gases can be measured simultaneously with another instrument to produce a compensation factor. In this case, the sensors were mounted inside a cell that was connected to the atmosphere with a tube whose inlet was close to that of the reference instrument. While the superiority of this technique was not a primary conclusion of the study[91], it showed the feasibility of calibrating outdoors for gas sensors.

# **3.3** Outdoor calibration

Co-location is the practice of placing a sensor next to a reference, running it for a time and then comparing the two datasets.

The SOGS-MOS sensors were calibrated by calculating the best fit to the reference instruments in the DEFRA AURN station on Leicester University campus[148], operated by Bureau Veritas. The NOx analyser is a Thermo-Scientific



FIGURE 3.10: Isometric diagram of the relative positioning of Zephyr sensors near the roof of the AURN station. The lip of the roof is roughly 3 meters from the ground.

model 42i, which is a chemiluminescence monitor capable of reporting NO, NO<sub>2</sub> and NOx concentrations, switching between NO and NO<sub>2</sub> modes in the manner described in section 1.5.4 on page 38. It has an adjustable range, but for NO<sub>2</sub> at  $955 \,\mu g \, m^{-3}$  it is capable of a precision of  $0.76 \,\mu g \, m^{-3}$  which is also the lower detectable limit. Measuring every ten seconds, it has a 40 second response time.

The ozone monitor is a Thermo-Scientific model 49i. This UV photometer has a precision of  $2 \,\mu g \, m^{-3}$  and a lower detectable limit of  $1 \,\mu g \, m^{-3}$ . The instrument response time is 20 seconds, with a 10 second lag.

The intake to the AURN station has an isokinetic pump that draws air down to the DEFRA reference instruments. The Zephyr prototypes were mounted on the exterior of the weather station building, as shown in figure 3.10. Aside from the reference instruments, there are two types of instrument involved in these experiments:

- A "long-term" SOGS-based experiment (LTE) using MOS sensors. There were two long term instruments the first used a version 0.2.5 SOGS base board and a single set of MOS sensors. While an older design, the signal processing and support electronics were similar to those of later versions used in this thesis. This LTE instrument was deployed at the AURN station from December 2014, before being replaced in early 2016. The second LTE, using a version 0.2.11 SOGS base board with a newer housing, is the one used in the analysis in this chapter. It ran for 15 months in total from February 2016 to June 2017.
- There were temporary mounting points installed for other SOGS-based MOS instruments, some of which were Zephyr prototypes. These were involved in several experiments but were all calibrated by being attached to plates along a horizontal crossbar next to the LTE as shown in figure 3.10. There are ten mounting points for these instruments, roughly 25cm apart, and all were used in the experiment described in section 6.1 on page 196. This group comprises any SOGS/Zephyr instrument mentioned in this thesis not specifically mentioned as being the long term experiment.

This calibration setup has issues with representivity, due to the distance between the sensor and the intake to the AURN's own instruments combined with the high spatial variability of the target gases[84], the effect of sunlight on the surrounding  $NO_2$  and  $O_3$  concentrations and the irregular flow of wind over the setup. The practical consequence of these issues is discussed in section 6 on page 195.

The AURN instruments take multiple measurements every minute, but these are averaged together into hourly intervals, with the timestamp of the resulting data matching the beginning of the hour over which the individual measurements are averaged[10]. For calibration, measurements taken from the Zephyr prototype instruments and the LTE must be averaged in the same way.

## 3.3.1 Physical configuration of the AURN site

The AURN station is located in a brick building on the University of Leicester campus, next to a leisure center and about 10 meters from University Road[148]. The AURN station and its environs are depicted in figure 3.11.



FIGURE 3.11: Map of the surroundings of the AURN station at the University of Leicester, with the station building and adjacent leisure center labelled in blue. North is directly up in the image.

Welford road is a busy thoroughfare that runs from the center of town south past the Leicester Royal Infirmary. University road is less busy except during rush hours. An experiment partly intending to measure the effects of these roads is described in section 6.2 on page 209. The leisure center contains a swimming pool, the chlorine from which might affect the sensors in unknown ways when the wind is blowing in certain directions. More information on this issue can be found in section 4.4 on page 166.

The anemometer for measuring AURN wind speeds is mounted on the roof of the building and operated by the university. There is a separate modelled wind speed for the AURN compiled by DEFRA that was not used in these experiments.

#### 3.3.2 The long-term experiment

The long-term experiment (LTE) provided data for fifteen months, albeit with some gaps during maintenance of the DEFRA reference instruments and technical issues with the SOGS prototype (version 0.2.3, see section A.2.2 on page 235). The LTE sensor is fitted with four MOS sensors of each type, with sensors of the same kind mounted about five centimeters apart. In the subsequent analysis, these sets of sensors are referred to as a, b, c and d. Sensor sets a and b are located on opposite sides of the same PCB, likewise with c and d. As mentioned in the previous chapter, the two sensors most relevant to our work are the oxidising sensor in MICS-4514, referred to as "OX", and the ozone sensor in MICS-2614, which is referred to as "O3". Consequently "O3a" refers to the MICS-2614 sensor belonging to set a.

To enhance reliability, the sensor was connected via USB to a laptop that recorded the sensor data in real time, but the sensor also recorded data to its onboard SD card. In spite of this, technical issues caused lengthy gaps in the reported data and also had an influence on the data received, as will be discussed below. Figure 3.12 shows the full long term experiment dataset. The data, at hourly intervals, covers 67.7% of the experiment's duration.



FIGURE 3.12: Data from the LTE dataset after zero voltage readings have been removed, showing typical voltage outputs from the OXa and O3a sensors alongside temperature (Temp.) and relative humidity (%RH) information, and the NO<sub>2</sub> and O<sub>3</sub> readings of the AURN reference instruments.

Occasional digital glitches (such as reports of zero resistance from the sensor ADCs) can result in unphysical data points that are major outliers - beyond ten times the standard deviation. Power supply issues could cause this to happen. To make analysis easier, these are omitted from the LTE dataset in the following.

#### 3.3.3 Raw sensor data

As discussed earlier, the sensors mounted outside on the AURN station will respond not just to their target gas, but also to environmental influences (such as temperature and humidity), and consequently the relation between their output voltage might not match that given in the datasheet, which is valid at 25°C and 50% relative humidity. By trimming the LTE data to include only data points where the temperature and humidity are close to the datasheet conditions, the extent of the interference of unknown gases in the outdoor environment interfere can be determined.

Figure 3.13 shows two scatter plots taken from a sample of the long-term experiment data, with each point selected on condition that relative humidity was between 45 and 55%, and the temperature between 23 and 27°C. The left plot of NO<sub>2</sub> shows very poor correlation; the O<sub>3</sub> plot is somewhat better, suggesting that the ozone sensor is responsive primarily to the ozone concentration.



FIGURE 3.13: Scatter plots of the resistance response of the oxidising (OX) and ozone (O3) sensors to target gases, for data taken within 45 and 55% relative humidity and 23 and 27°C. Note how, in the left panel, the NO<sub>2</sub> data for sensors c and d occur in two distinct domains, offset vertically by one to one and a half ohm.

The signals from O3 sensor a and OX sensor d occur in two very distinct patches in figure 3.13. A scatter plot shows the extent of the correlation between two variables, and the shape of the patch in the scatter plot can provide clues as to the exact relation between them. Two well defined patches, however, imply there are two different relations that govern the correlation between these sensor outputs. This could be a change that occurs after a specific time, and it could be a change that occurs to one sensor, or all of them. Critically, the higher-resistance patch of OX sensor d shows no response at all to changing reference  $NO_2$  levels. To investigate the relation between these sensors over time, a graph of changing correlation over time can be plotted. This graph shows the relation of a specific sensor to the average of all of the sensors, and is thus independent of the gas concentration response function:

$$M_i = \frac{V_i}{\bar{V}} \tag{3.14}$$

Where M is the "correlation meander", V is a sensor voltage,  $\overline{V}$  is the average of all sensor voltages and i identifies the particular sensor (a, b, c or d). Plotting the correlation meander can make explicit the time component of a change in relation between the two patches in figure 3.13. If the form of the meander changes suddenly at a certain point in time, it means that at the very least calibrations performed before that change will not be valid for data afterward. The correlation meander of the LTE timeseries is shown in figure 3.14.

For sensors that are responding purely to changes in atmospheric composition, the correlation over time would be perfectly flat. However, there are several large transitions in the period before May 2016, during which the relationship between the sensors is highly variable. There is another large transition in the shape of the meander toward mid-November 2016.

These two periods correspond with power supply issues on the AURN station. The power supply for a set of sensors on the AURN station was not adequate for large numbers of sensors before May 2016, and was often overloaded resulting in multiple undervoltages of the sensor's auxiliary power supply. The effect this has on the sensors is significant and sustained. However the largest transition between sensors is during the period marked on figure 3.14 with a red dashed line - a power surge at the AURN station's main power supply occurred on the 15th of November,



FIGURE 3.14: Correlation meander over time for long term experiment sensors, both O3 and OX. Two gaps that indicate obvious changes in correlation are marked with dashed red lines, and have been labelled with possible explanations given in the text. The y units are dimensionless, given that they are the ratio between one variable and the average of several of the same kind. The meander's value is very close to one for most of the timeseries, varying by less than a thousandth.

and seems to have severely affected the sensors. Another power cut occurred on the 12th of September, which affected other sensors that were calibrating on the AURN station, but not the LTE.

The latter event might have been significant for the instrument because the laptop's USB connection power would have kept the instrument's MCU alive with the laptop battery, and it would have scheduled taking readings, even as the sensor was unable to save data to the SD card which depends on the auxiliary supply connected to the mains of the AURN station. Many of the long gaps that occurred before this time were due to the fact that the sensor firmware would shut off the SD card saving after a single error, and consequently any interruption in the LTE power while it was connected to the USB cable would stop it saving any data. This error was worked around until it was fixed in a firmware update on the 18th of January 2017. However, during maintenance throughout the year the sensor would have been disconnected for up to an hour, and in no case was the change in correlation as severe as during the power cut on the 15th of November 2016 and in the "brownout period" before May of that year.


(b) Colour-coded timeseries of raw sensor resistance

FIGURE 3.15: (a) Scatter plots showing relation between a, b, c and d sensor resistances for both OX (upper right graphs) and O3 (lower left). The scale is the same on every plot of the same type of sensor: 1783 to 1786 ohms for the OX sensors and 1781 to 1786 for the O3 sensors. The scale is less important than the shape of the different coloured patches, and the colours correspond to different time periods, shown in (b). These time period colours are delineated by the end of the brownouts in May 2016 and then the power cuts in November 2016.

A set of scatter plots for the oxidising gas and ozone sensor resistances for sensors a-d is shown in figure 3.15(a). This plot shows the changes in relative sensor response, identifying changes after the power supply events.



FIGURE 3.16: Scatter showing resistance of sensor O3b vs O3d during the long term experiment, colour coded in the same way as 3.15(b).

Figure 3.16 is particularly clear in illustrating the effects of unreliable power, as the date from the brownout period is split into several patches. The patch after the power surge is mostly intact, as is the patch from the middle period. In figure 3.15(a) the split between OX sensor correlation patches is even more pronounced. The post-surge period's patches have horizontal bars that indicate the sensor is insensitive at some points; while the patch for the middle period is better on this count, there is also some splitting in this case in every one of the OX sensor comparisons.

The fact that the disturbances in correlation occurred at the same time as electrical issues suggests that the sensors might not be comparable across these interrupted periods. While serious interruptions and strains on the sensor power supply might have affected the devices in unpredictable ways, the other data might yet be consistent and it will take further analysis in section 3.4.4 on page 127 to determine the most reliable time period.

# **3.4** Preliminary analysis

This section describes the two important practical issues that must be resolved before the sensors can be calibrated. One is choosing the right fitting equation, which relates the raw sensor output discussed in the previous section to the actual gas concentrations. Several of these are proposed in section 3.4.1 on the next page and tested in section 3.5 on page 130. The other is determining the best time duration for calibrations, discussed in section 3.4.2 on page 120.

A custom set of data analysis tools were developed for use with SOGS, based on the NumPy python package[223]. The architecture for this software revolved around "analysis objects" which contained fields corresponding to a dataset, and methods for manipulating and processing that data. These objects simplify manipulation of large datasets, and they also track operations applied to the data to make debugging easier. This methodology is based on the ideas of the LISA Technology Package Data Analysis system developed for the eLISA (Laser Interferometer Space Antenna) spacecraft program[224][225], although the codebase is not descended from it. In addition to streamlining the handling of large data sets, the analysis software was able to easily visualize sensor and housekeeping data. Fit objects were developed that allowed different methodologies for fitting data to be applied to the analysis object timeseries in an efficient way.

#### 3.4.1 Candidate equations

The simple linear equations (3.9 and 3.10) were adequate for tightly controlled conditions, but they contain no correction factors for other gases, or humidity and temperature, all of which are known to be an issue for these sensors in realistic environments. In other work it has been demonstrated that NO<sub>2</sub> MOS sensors have a significant, and correctable, cross-sensitivity with O<sub>3</sub>[170]. Several equations that incorporate this interference, in addition to possible temperature and humidity effects, have been developed empirically. The V terms in these equations refer to voltages, which are proportional to the sensor resistance plus a constant. The two voltages are  $V_{OX}$ , the voltage for the oxidising gas sensor, and  $V_{O3}$ , the voltage from the ozone sensor. RH and T are the outputs from the relative humidity and temperature sensor respectively. The k terms are constants to be found during calibration. For NO<sub>2</sub>, the equations all have the form of a linear sum of terms:

• *Linear* equation:

$$[NO_2] = k_c + k_1 V_{OX} + k_2 V_{OX} RH + k_3 V_{OX} T + k_4 V_{O3} + k_5 V_{O3} RH + k_6 V_{O3} T$$
(3.15)

This equation assumes that the influence of temperature and humidity is proportional to total sensor resistance. Both  $O_3$  and  $NO_2$  have the same form of influence, in this case that of a linear offset to the baseline concentration  $k_c$ . Temperature changes affect the strength of the influence of each gas, to an extend governed by the  $k_3$  and  $k_6$  terms. Humidity functions identically, through the  $k_2$  and  $k_5$  terms. • *Multiplicative linear* equation:

$$[NO_2] = k_c + k_1 V_{OX} + k_2 V_{OX} RH + k_3 V_{OX} T + k_4 V_{OX} V_{O3} RH + k_5 V_{OX} V_{O3} T$$
(3.16)

This equation assumes both the temperature and humidity and, unlike the Linear equation, the ozone interference, are all proportional to the sensor resistance of the oxidising gas sensor. The oxidising gas sensor is a scalar factor multiplying the effects of the ozone interference through the  $k_4$ ,  $k_5$  and  $k_6$  terms. The other terms function identically to the linear equation.

• Inverse NO<sub>2</sub> Linear O<sub>3</sub> multiplicative equation:

$$[NO_2] = k_c + \frac{k_1}{V_{OX}} + \frac{k_2}{V_{OX}}RH + \frac{k_3}{V_{OX}}T + \frac{k_4}{V_{OX}}V_{O3}RH + \frac{k_5}{V_{OX}}V_{O3}T \quad (3.17)$$

Similar to the Multiplicative linear equation, except using inverse terms for the oxidising sensor voltage, which may more closely match the non-linear response of the sensors as discussed regarding cell calibration in section 3.2.2 on page 99.

• Inverse NO2 equation: This equation uses inverse terms for both  $V_{OX}$  and  $V_{O3}$ . Abbreviated as Inverse.

$$[NO_2] = k_c + \frac{k_1}{V_{OX}} + \frac{k_2}{V_{OX}}RH + \frac{k_3}{V_{OX}}T + \frac{k_2}{V_{OX}V_{O3}}RH + \frac{k_2}{V_{OX}V_{O3}}T \quad (3.18)$$

These calibration equations for  $NO_2$  account for temperature, relative humidity and ozone interference in the sensor response.

The equations for ozone are simpler, as they assume no interference from oxidising gases with the O3 sensor operation:

• *Linear* equation:

$$[O_3] = k_c + k_1 V_{O3} + k_2 V_{O3} R H + k_3 V_{O3} T$$
(3.19)

This simple equation keeps the multiplicative terms for temperature  $(k_3)$  and humidity  $(k_2)$ , producing an influence proportional to the sensor voltage.

• Inverse equation:

$$[O_3] = k_c + \frac{k_1}{V_{O3}} + \frac{k_2}{V_{O3}}RH + \frac{k_3}{V_{O3}}T$$
(3.20)

An equation for Ozone with inverted voltage terms.

• k+Inverse equation:

$$[O_3] = k_c + \frac{k_1}{k_4 + V_{O3}} + \frac{k_2 R H}{k_4 + V_{O3}} + \frac{k_3 T}{k_4 + V_{O3}}$$
(3.21)

This modification of the inverse equation has an additional term  $(k_4)$  in the denominator of each to "soften" the curve, in case this helps match the non-linear response of the sensor.

These equations are largely designed to approximate the response of the sensors observed during early testing in the closed gas cell (see section 3.2.2 on page 99). To determine which equation works best in practise, first the optimal calibration period must be found (section 3.4.2 on the following page), and a methodology must be developed for evaluating the equations in a systematic way. While the constants discovered during fitting will mean that each equation's output will be similar during the calibration period, changes in the raw sensor data will affect each equation differently - for example, when the voltage from the oxidising gas sensor is low, the output of the linear equation 3.15 will respond less strongly to changes in the ozone sensor voltage than the inverse equation 3.18. Consequently the equation outputs will diverge over the long term experiment as a consequence of the conditions during the calibration period. Section 3.4.3 on page 124 describes a method that can allow the quality of the predictions produced by the various equations to be compared, incorporating the effects of calibrations taken over the entire dataset rather than a single period. This method will also answer the questions raised in section 3.3.3 on page 112 regarding the integrity of different sections of the LTE dataset.

Despite the electronic issues with the sensors, it is clear that they respond to varying gas concentrations. For example, using equation 3.16 for NO<sub>2</sub> and equation 3.20 for O<sub>3</sub> fitting on a month-long dataset produces the timeseries shown in figure 3.17. Throughout the subsequent sections it will be noted where assumptions have been made about the best fitting equation to use.



FIGURE 3.17: A month-long timeseries from the long term experiment between 17th of March and 17th of April 2016, with predictions from the MOS sensor data using equations 3.16 and 3.20 for NO<sub>2</sub> and O<sub>3</sub> respectively. The black lines in the NO<sub>2</sub> and O<sub>3</sub> sections are reference data from the AURN station instruments. The week-long calibration period is shown as a red box on the humidity timeseries.

#### **3.4.2** Optimal calibration time

The optimal length of time for calibrating a sensor is an important question. Beyond the practical consideration that a short calibration time is more convenient, the system being measured also imposes constraints on experiments. The level of  $NO_2$  in the atmosphere varies on a scale of seconds and generally follows diurnal cycle, but concentrations of interfering gases may vary on a daily or even a seasonal scale[226]. If the calibration is done over a long period, the unmeasured background gases could vary significantly enough to offset the sensor response. With no way of compensating for this effect, the accuracy of the fit predictions would be reduced.

For this analysis, calibrations are repeatedly taken across the entire length of the LTE timeseries, and the quality of the fit is calculated for the week immediately afterward - provided there is contiguous data within this time. An arbitrary equation can be made to fit, or possibly overfit, the data, if the calibration and validation period are the same. To avoid this trap, fits must be evaluated after the calibration period is over[227].



FIGURE 3.18: Various metrics of quality of calibration vs. length of calibration period, for all four sets of sensors. NO<sub>2</sub> and O<sub>3</sub> are fitted using the "inverse NO2 linear O3" equation 3.17 and the "inverse" equation 3.20 respectively. The different coloured lines correspond to OX and O3 sensors a, b, c, and d, and the best point on each quality curve is marked with a circle. The metrics of quality compare the predictions of the sensors with those of the AURN station during the validation period, a week directly after the calibration.

Figure 3.18 is a plot of the average quality of sensor fits as a function of the length of the calibration period, for NO<sub>2</sub> and O<sub>3</sub>. The average of these quality metrics are shown for each calibration length on the y-axes of the figure.

In figure 3.18, for all the quality metrics, very short fits of less than three or so days perform the worst, but for longer periods the quality tends to level out, or even decline slightly. This is very clear in the NO<sub>2</sub> fit fractional errors and *RSE* (panel a). For O<sub>3</sub>, *PCC* is fairly flat with time (panel f), with the best results between 4 and 7 days, but for the other metrics (in panel d and e) the fits very gradually improve the longer the calibration period. For NO<sub>2</sub> it is a different story in sensors c and d especially. The phenomenon of fits becoming less accurate the longer the calibration is taken for might be due to slowly varying background gas interference. The becomes more significant the longer the period of calibration.

Degradation in the sensors during the calibration period itself must be mentioned. The resistance of the heater element will decrease over time as the sensor is used [203]. In this instrument (and most of the circuits that make use of MOS sensors) the heater is driven by a constant voltage supply. This means the temperature of the sensor element will gradually increase over time, changing its sensitivity to different gases in the atmosphere. The sensor element itself can experience chemical changes that affect its responsiveness [228]. While significant in explaining the degradation of the sensor performance in the long term, these phenomena affect the sensor differently to interfering gases. The concentrations of these gases might vary in an unpredictable way, even returning to previous levels, during the course of an experiment. As stated in the previous paragraph, these random fluctuations cause longer calibration periods to produce less accurate predictions. The sensor's degradation over time is a predictable one-way process. During the time when the sensor must make useful predictions, the element is degrading at the same rate, regardless of the length of calibration period. Therefore changes in calibration length might help mitigate the effects of varying background gas concentrations, but will not help with the effects of sensor degradation.

A week-long calibration period seems to be optimal for getting good quality fits with either target gas. However, leaving the sensor for longer has very minor benefits.

For the purposes of deployment in the field, the sensors are calibrated at the AURN station for a period of two weeks. The first week being the source of the calibration fit, and the other being a period over which the fit can be evaluated to see if the sensor's output has the desired accuracy (the metrics of calibration quality are investigated more fully in section 3.7 on page 137). The choice of time period is backed up by the analysis in this section (and figure 3.18). This calibration period is also practically convenient. Additionally, with hourly AURN measurements, a week-long calibration yields a sequence of data with enough points for a fit to be reliably found. For most purposes in this thesis the calibration is followed by a week-long validation period, unless explicitly stated otherwise (for example in section 4.2.1 on page 150).

#### 3.4.3 Combined long-term data analysis

A technique was developed to show how effective a particular calibration equation, or method of analysis, was for predicting data at an arbitrary time from calibration. This technique involves generalizing the deployment calibration procedure (illustrated in figure 3.19), which takes place over two consecutive weeks, toward an entire timeseries measured in years. It requires a dataset where an instrument under test (the LTE) is operated alongside a reference instrument (in the AURN station, see section 3.3 on page 105) for a long period of time.

This Combined Analysis For Predictions Of Long Datasets (CAFPOLD<sup>1</sup>) is a broadening of the scope of the simple procedure of calibration-prediction-evaluation shown in figure 3.19, with the aim of making it general to an entire dataset.



FIGURE 3.19: Visual representation of a calibration taken over a weeklong part of a larger data series (dark orange) being used to create a weeklong prediction (light orange), from which quality metrics can be taken by comparing the prediction with the reference instrument. The grey line represents the underlying data being produced by both the sensor being calibrated and the reference.

The procedure for CAFPOLD analysis has six steps:

1. The AURN reference and LTE sensor timeseries are divided into "calibration sections" between 6 and 7 days long when data was available. This contracting to a minimum of 6 days where there were gaps in the data. The divisions are made at hourly intervals to match the averaging period of the reference data.

<sup>&</sup>lt;sup>1</sup>Better names, and hence acronyms, for this data analysis technique may be submitted for consideration to the author's email address.

2. For each of these sections, a linear regression fit is taken and predictions made for the period of the whole timeseries. Occasionally the fit algorithm would fail to resolve the coefficients for the equation in question, in which case the fit is ignored and the next section is processed. If the fit successfully resolves, it produces a prediction over the entire dataset from the fitted coefficients:



FIGURE 3.20: Visual representation of set of predictions over the whole reference timeseries. t=0 at the beginning of the particular calibration section that produced the predictions.

3. Ratings (statistics measuring goodness of fit) are calculated for each prediction. The statistics that will be used are defined in section 2.1 on page 51: the Pearson correlation coefficient between prediction and reference (equation 2.3), residual standard error (equation 2.1) and residual fractional error (equation 2.2).



FIGURE 3.21: The grey hexagons represent ratings, the darker the hexagon the better the rating. Effectively the closer the rating is to the calibration period, the better it is likely to be.

4. This analytical process is repeated for every calibration period throughout the dataset. Each repetition produces an array of ratings vs. time from calibration, for the fit is being judged by those ratings.



FIGURE 3.22: Collection of ratings arrays for different calibration sections. This concept is further illustrated in figure 3.25.

5. The timestamps of the ratings arrays are shifted so that they are aligned at t=0.



FIGURE 3.23: Shifting the ratings arrays to align their timestamps.

6. Finally, the average of all ratings are taken in day-long bins. This produces a single timeseries, showing the likely quality of a fit for a given time from calibration:



FIGURE 3.24: Averaging the ratings together. The red curve is an example of a series of averaged ratings.

Taking the standard deviation of the ratings in each day-bin gives the level of variation in the quality of calibrations, for the time from calibration that the particular bin represents. For example, a higher standard deviation means that some calibrations were far better than others for predicting data at that point.

The output of the CAFPOLD analysis may be plotted as a timeseries of the various quality metrics, with respect to time from calibration. Figure 3.25 is a guide to interpreting these graphs.



FIGURE 3.25: Explanation of the time axis of CAFPOLD plots. The "source" axis shows arbitrary predicted and reference gas concentration data. The red box shows the section of the source data used for fitting a calibration equation. The "CAFPOLD" axis is the graph of a given quality metric over time. The purple rectangles on the source axis show the time periods over which the quality metric is calculated. The value of that metric corresponds to a point on the CAFPOLD axis indicated by a circle. Point  $\mathbf{A}$  is on negative time because it represents a prediction made from calibration data that occurred in the future. The blue lines at  $t_c$  on the CAFPOLD graph bracket the region of points where the quality metrics are calculated for predictions partially during calibration. For example the first half of the rating for point **B** covers data just before calibration, the latter half covers some data from within the calibration period. Point C shows the point on the CAFPOLD graph corresponding to quality metrics calculated directly after the calibration period ended. It should be noted that for this graph a single calibration on a single section of reference data is displayed. A real CAFPOLD analysis takes the average of many such calibrations.

#### 3.4.4 Issues with long term experiment data

A CAFPOLD analysis was taken during the three different periods of data discussed above in section 3.3.3 on page 111. The three sections of data are labelled:

- Zeroth, the section between February 2016 and the end of May of the same year, ie. the period during which there were power brownouts.
- *First*, the section from June to mid November 2016, during which the station was functioning normally.
- *Second*, the section after November 15th 2016, which occurred after the power cut at the AURN station.

To make it easier to compare the three sections, the ratings for the four sensors from each section were averaged together. The average uses the absolute value for time, meaning in total eight data points are averaged to produce a single output datum. The result is shown in figure 3.26, and quantities are given in table 3.2. This analysis used the "Inverse" equations for fitting NO<sub>2</sub> (equation 3.18) and O<sub>3</sub> (equation 3.20). The other fitting equations from section 3.4.1 on page 117 produced a graph with similar appearance. For *RSE* and *FE*, a lower number implies better performance; for *PCC*, higher is better.

			$ m RSE~\mu gm^{-3}$			FE			PCC	2
	Section		Typical	SD		Typical	SD		Typical	SD
NO2	Zeroth	a)	263.7	573.2		13.9	30.2		0.52	0.11
	First		26.4	6.74	b)	1.2	0.95	c)	0.54	0.11
	Second		32.3	6.32		0.5	0.15		0.59	0.22
	Zeroth		40.8	9.27		1.4	0.92		0.64	0.07
O3	First	d)	25.0	5.62	e)	1.3	1.02	f)	0.81	0.06
	Second		24.7	4.71		2.6	1.64		0.74	0.07

TABLE 3.2: Performance statistics for analyses of all three sections. The "typical" values are averages of those statistics across the CAFPOLD timeseries and all four of the sensors in the Long Term Experiment. The SD (standard deviation) values given are the squared sum of the residuals from the timeseries average and the inter-sensor average. The relevant subfigures from figure 3.26 is shown by the bracketed letters.

The extremely high RSE and FE for the zeroth section (a full order of magnitude higher than the first section for NO<sub>2</sub>) is a practical demonstration of the effects the unreliable power supply had on the sensor equipment. The *PCC* for the zeroth



FIGURE 3.26: CAFPOLD analysis comparing different sections of LTE data. The colour scheme is the same for NO<sub>2</sub> and O<sub>3</sub>. The start of the calibration period is at zero on the X axis. The end of the calibration, which lasts a week, is marked by the vertical blue line. Each data point represents an averaged rating, and the X axis position of a point marks the start of the period within which the rating is computed. Plots a) b) and c) give the ratings for NO<sub>2</sub>, and d) e) and f) for O<sub>3</sub>, with one graph for each metric of quality. Gaps in the plot traces correspond to missing data or sections where the fit could not produce a prediction. The black horizontal lines represent different thresholds on the *RSE* graphs a) and d): 20 µg m<sup>-3</sup> for the solid line, 40 µg m<sup>-3</sup> for the dashed line.

section is worse than the other sections, which tells us that between brownout events the calibrated sensor outputs still followed the target gas concentration. The deficiency of this rating is far less dramatic, however.

The distinction between the first and second sections is more subtle. Generally the first section has marginally better RSE than the second section for NO<sub>2</sub> and similar RSE in subfigure d) for O<sub>3</sub> - within one standard deviation. The fractional error is comparable for NO<sub>2</sub> in b). For ozone, FE is much better for the first section than the second in as shown in panel e). PCC for NO<sub>2</sub> is roughly equivalent for both sections in panel c), and somewhat better for O<sub>3</sub> in the first section as shown in f). Overall the first section has lower error than the second (although the difference is sometimes marginal). Given the prospect of damage to the sensors from the more severe power supply issues in November 2016 only the first section will be used for further analysis.

# 3.5 Comparison of fitting equations

A similar CAFPOLD analysis to that in the previous section was carried out to determine the usefulness of the different equations described in section 3.4.1 on page 117, evaluating them for the "first" section of data between June 1st and November the 14th. Once again the results for the four individual sensors were averaged together and the absolute time from calibration was used. The results are shown in figure 3.27, with a table of values given in table 3.3.

				RS	Е		FE			PCC		
	Equation			Typical	SD		Typical	SD		Typical	SD	
	Linear*	3.15		97.7	138.9		2.41	2.7		0.38	0.17	
NO <sub>2</sub>	Inverse	3.18	a)	26.4	6.74	b)	1.24	0.95		0.54	0.11	
	Multip.*	3.16		60.4	96.52	0)	1.55	1.70	6)	0.49	0.12	
	Inv. Lin.	3.17		29.5	8.3		1.32	1.18		0.50	0.12	
	Linear	3.19		32.0	9.09		2.10	2.08		0.74	0.06	
$O_3$	Inverse	3.20	(d)	25.0	5.62	e)	1.26	1.02	f)	0.81	0.06	
	k+Inverse	3.21		36.1	31.40		1.01	0.83		0.82	0.08	

TABLE 3.3: Table of results from the analysis of different equation performance. The SD ratings are the combined residuals from the timeseries average and the average of the sensors a, b, c and d. The linear and multiplicative equations (marked with a \*) for NO<sub>2</sub> sensor b have been omitted from the statistics here. The references to the left of each block of figures is to the graphs in figure 3.27, and the equation numbers are listed next to the equation heading.

The graph in figure 3.27 shows spikes in all three metrics for the linear and multiplicative equations predicting  $NO_2$ , in panels a) b) and c). This is down to sensor b, for which the equations were unable to resolve finite results for roughly half the time. In table 3.3, the data from sensor b have been omitted from the averages - the rest of the sensors resolved normally, and these results are still useful for comparison.



FIGURE 3.27: CAFPOLD analysis comparing different fitting equations. The vertical blue line signifies a week from calibration, which is at zero on the X axis. There are three plots for each gas for the three quality metrics. Gaps in the plot traces correspond to missing data or sections where the fit could not produce a prediction. Once again the specification thresholds are denoted with black horizontal lines,  $20 \,\mu g \,m^{-3}$  for the solid line,  $40 \,\mu g \,m^{-3}$  for the dashed line.

#### 3.5.1 Discussion

For NO<sub>2</sub>, the Inverse equation (3.18) has the best statistics for this data accross all three metrics. Average fractional error, at 1.24, is still unacceptably high for most of the data series though, although figure 3.27 b) shows that for all equations the fractional error is less than one for the first month after calibration and becomes significantly worse after two months. The pure Linear equation 3.15 performed very poorly, with residual standard error four times that of the Inverse equation, and particularly poor *PCC* statistics even a week after calibration. The hybrid Inverse NO<sub>2</sub>-Linear O<sub>3</sub> equation 3.17 was second in performance to equation 3.18, and the multiplicative equation 3.16 performed comparably to the best two close to calibration, but was significantly worse further from calibration time. The comparatively superior performance of equations involving an inversely proportional response to oxidising gas sensor voltage implies that the response of the sensors in this region more closely matches the curve of an inverse function rather than a directly proportional one. The very inferior predictions of the simple linear equation imply further that the interference from ozone in the oxidising gas sensor becomes more significant when the oxidising gas sensor output is higher. Of the four, the best for general use is the Inverse equation 3.18.

The situation with ozone is more complex. While the simple linear 3.19 equation was inferior, particularly for PCC and FE, it was nowhere near as bad as the linear equation for NO<sub>2</sub>. The simple inverse O<sub>3</sub> equation 3.20 predictions gave the best RSE, but were worse in terms of fractional error than the k+inverse equation 3.21. However the latter equation gave predictions with not just a higher RSE, but a highly variable one. Figure 3.27 d) shows why - for the most part the equation's predictions were similar to the inverse O<sub>3</sub> equation 3.20, and even better within a month of the calibration time, but there is a period between 2.7 and 3.7 months from calibration in where the RSE of the k+inverse predictions suddenly skyrockets. There is also a smaller blip in the k+inverse RSE performance three weeks from calibration. The unreliability of this equation means that in spite of its superior fractional error performance, it is not suitable for general use and the best equation for O<sub>3</sub> with these sensors is the inverse equation 3.20.

# 3.6 Calibrated sensor performance

With the best fitting equations for both gases determined and a relatively error-free set of data, it is possible to perform a detailed CAFPOLD analysis on the sensor data to examine their individual performance in this calibration environment. In this analysis a "baseline" data series was included, with the MOS sensor data "zeroed out", with all voltages set to one but humidity and temperature data left untouched. This provides a comparison with a completely unresponsive sensor, which is a useful diagnostic for devices where the airflow to the sensor has stagnated or that have experienced an electrical fault. The results are shown in figure 3.28, and performance numbers are given in table 3.4 on the next page.



FIGURE 3.28: CAFPOLD residual standard error RSE, logarithmic fractional error FE and fit correlation PCC against time from calibration tFC for individual adjacent sensor elements, for NO<sub>2</sub> (graphs a-c) and O<sub>3</sub> (graphs d-f). The black horizontal lines are the two target sensitivity levels for RSE, and at zero for PCC. The coloured lines represent the different sensor elements, whereas the grey line is the baseline measurement. The shaded area of subplot a) and d) show the range of the standard deviation of the instrument RSE, within which 63% of the prediction data series lies.

To obtain figures for the change in performance over time, the prediction from all four sensors as shown in figure 3.28 was averaged together, and the absolute value for the time from calibration was used. The linear regression for the resulting data is the prediction of the sensor's performance at day zero (ie the calibration and prediction test periods are over the same time), and more importantly the change per day of each performance metric.

or		Residual standard error			Fractional error				Pearson Correlation				To spec		
	ens	1 Month		5 Month		1 Month		5 Month		1 Month		5 Month		1	5
	Ś	RSE	SD	RSE	SD	FE	SD	FE	SD	PCC	SD	PCC	SD		
	a	23.2	3.24	27.2	5.61	0.77	0.40	1.67	1.39	0.65	0.04	0.54	0.11	17%	5%
NO2	b	23.1	3.32	25.6	3.85	0.69	0.27	1.26	0.82	0.62	0.06	0.54	0.10	17%	6%
NUZ	С	23.4	3.64	25.4	6.78	0.63	0.25	1.14	0.79	0.62	0.05	0.56	0.11	12%	12%
	d	24.9	5.38	27.5	10.4	0.58	0.16	0.91	0.58	0.60	0.06	0.54	0.12	10%	12%
	a	23.8	4.02	25.2	5.59	1.25	0.80	1.45	1.03	0.80	0.04	0.81	0.06	20%	13%
O3	b	23.3	3.88	24.6	4.92	0.94	0.50	1.06	0.74	0.79	0.04	0.79	0.07	30%	19%
	С	22.4	3.72	24.9	6.04	0.99	0.55	1.19	1.02	0.81	0.04	0.82	0.06	35%	19%
	d	23.5	4.04	25.4	5.90	1.22	0.74	1.32	1.22	0.80	0.05	0.81	0.06	27%	15%

TABLE 3.4: Summary of sensor performance statistics of all four sensors, for both  $NO_2$  and  $O_3$ . The colour grading indicates relative quality, with the darkest green indicating the best performance figure relative to the other sensors, and darker yellow means a higher standard deviation for the performance statistics, and hence more inconsistent predictions.

		$ m RSE~\mu gm^{-3}$	FE	PCC
$NO_2$	Day zero	$20.91\pm0.88$	$0.54 \pm 0.086$	$0.60 \pm 0.020$
	Change/day	$0.079 \pm 0.011$	$0.010 \pm 0.0010$	$-0.00075 \pm 0.00024$
$\cap$	Day	$22.57 \pm 1.56$	$0.61 \pm 0.16$	$0.82 \pm 0.015$
$O_3$	Change/day	$0.034 \pm 0.019$	$0.010 \pm 0.020$	$-0.00011 \pm 0.00018$

TABLE 3.5: Linear regression showing the day zero and change per day statistics for RSE, FE and PCC from the CAFPOLD analysis of the LTE predictions using the Inverse equation for NO<sub>2</sub> and O<sub>3</sub>. The uncertainties listed are the 95% confidence interval for the linear regression.

#### 3.6.1 Discussion

Figure 3.28 a) and d) shows RSE vs. time from calibration some features in common for all four of the proximate sensors for NO<sub>2</sub> and O<sub>3</sub> respectively. The large instability of RSE in earlier time periods in the a) graph indicates that fits taken toward the end of the section were somehow not suited for measurements at the beginning, as all four sensors perform worse than the baseline. For sensors c and d the RSE in NO<sub>2</sub> is substantially higher in the early period, but in figure 3.28 b) it is clear that the factional error of these two sensors is markedly lower.

Sensors c and d were spatially close together inside the Zephyr instrument, but

not directly on the same board as sensors a and b. Their substantially worse RSE performance toward the beginning of the analysis timeseries in figure 3.28 a) indicates that calibrations taken at the end of the data do not make good predictions for these sensors at the beginning of the dataset. Sensor a also performed worse than the baseline during this part of the data, a serious flaw, although nowhere near as deficient as c and d. *PCC* performance for all the sensors was better than the baseline (which itself hovered around zero for the entire data series, as should be expected mathematically) during the entire data series as shown in in figure 3.28 c), so the sensors were at least responding to relative changes in gas concentration, but the prediction must have had an offset from the reference data of similar magnitude to the sensor's analysed *RSE* values in this period.

Within one month of calibration all four sensors achieved similar levels of performance in all metrics. Their dramatic increase in fractional error for c and d early in the data series has several explanations, some based on the sensor elements themselves, and some on the hardware driving them. The printed circuit board (PCB) for sensors c and d share a step-down power supply for their heating element, and it is possible this is warming them to a different temperature, which would change their sensitivity to atmospheric gases[197]. Another explanation is simple manufacturing differences between the sensors, although the same kind of manufacturing defect shared between each pair of sensors is suspicious. The location of the PCB for c and d within the casing might cause a significant difference in the quality of air reaching them, but this wouldn't explain the similar performance level close to calibration, nor the relatively consistent *PCC* values over the whole timeseries.

Naturally fits prediction generally have the best quality statistics during the week they were taken. The *PCC* recorded at day zero averages 0.74 for all sensors in  $NO_2$  during the calibration period, but decreases rapidly beyond a month as shown in 3.28 c). The fitted decline in *PCC* shown in table 3.5 on the preceding page is more gradual, and shows the expected peak given a linear decay as around 0.60. For  $O_3$ , the sensors perform on worse than the baseline toward +5 months in *RSE*, as can be seen on graph d). Sensors c and d both have higher residual error than the other sensors, and are again correlated with each other. Fractional error also exceeds 1.0 for these sensors during this period. However *PCC* is close to 0.8 in 3.28 f) for the ozone sensors, implying the sensors are still responding to  $O_3$  in the atmosphere, but are "drifting" away from the true values. This bears out other studies of these sensors which show inconsistent results - some good[177], some poor[50], particularly far from the calibration time[188].

Fits become less reliable further out, with the 3.28 a) and d) showing typical RSE comparable to the expected concentrations of the target gases in the atmosphere. The *FE* graphs b) and e) underline this fact; only in the weeks surrounding calibration is the sensor error consistently smaller than the magnitude of the reference instrument's reading for NO<sub>2</sub> and O<sub>3</sub>. The relatively high *PCC* for both gases throughout the dataset, which only very gradually changes, suggests that while the sensors are still responding to changes in target gas concentration they have become offset from the reference instruments over time. Still, this is clearly an unsatisfactory result.

In the table of sensor performance 3.4 on page 134, the mean values were obtained by averaging over both the whole timespan of the analysis (roughly 5 months), and the month long section closest to calibration. The mean of the CAFPOLD figures for each statistic (RSE, FE and PCC) is a general indication of sensor performance, and the standard deviation of these statistics shows how much the sensor performance can vary from these averages - so roughly 68% of the time, the sensor calibrations will produce results that fall within those of the mean.

The "To Spec" column in table 3.4 on page 134 shows the fraction of data points produced by each sensor that meet the design specification (standard deviation less than  $20 \,\mu g \, m^{-3}$ ) outlined in section 2.1.1 on page 53. While the sensors are capable of *RSE* that is usually less than a  $25 \,\mu g \, m^{-3}$  threshold, over five months

no sensor measuring either gas meets the design specification more than 20% of the time. Over one month the sensors did not meet specification more than 35% of the time for  $O_3$ , and for  $NO_2$  no more than 17%.

There are tantalizing hints in the data series and the above table that the unsatisfactory performance of these sensors, particularly beyond two months away from calibration, can be improved. The lowest end of the standard deviation region (shaded area on the *RSE* graph a) in figure 3.28) hovers at  $20 \,\mu g \,m^{-3}$  even when the mean of *RSE* is closer to twice that amount around -4 months, and this is the standard deviation of the quality of all the individual fits. Some fits therefore give much better predictions than others. If the quality of these fits is due to some environmental condition during calibration and they are consistently good throughout the year, it might be possible to test for these conditions and discard the poorer fits, an avenue which is now explored in section 3.7.

# 3.7 Selecting good fits

The predictions a fit of the LTE data can make varies greatly in quality depending on the exact days over which the calibration was made. Some fits produced predictions that were better than the specification, some far worse.

One way to improve the overall performance of the sensors is to use a validation period to judge the quality of the calibration, and discard that calibration if it does not pass a statistical "goodness" test. The validation period should be immediately after the calibration period. This section will test the hypothesis that a calibration which produces predictions with poor statistical quality during the validation period would likewise underperform over the rest of the data series. If this is true, then using this technique might potentially extend the time an instrument under test has to remain at the calibration site, in exchange for that instrument more accurately predicting NO<sub>2</sub> and O<sub>3</sub>. The three statistical metrics of quality discussed so far can be used to judge fits during the validation period. The measure of a good calibration is that one of these metrics passes a given threshold value during the validation period.

The choice of threshold and metric will be a tradeoff depending on the application. Setting a more stringent threshold might improve the prospects of good sensor performance, but it will also lead to more fits being rejected. The decision ultimately depends on the application and the needs of the sensor's end user.

A modification of the CAFPOLD analysis can be used to determine the performance effects of each type of threshold for the entire dataset. Referring to the six step process described in section 3.4.3 on page 124, the prediction is tested just after the original fit is made for the calibration section (step 2), and if it fails then it is discarded, in the same way it would be if the fitting algorithm was unable to resolve the equation coefficients. In some cases large numbers of predictions had to be discarded, and if at stage 6 there were less than five timeseries to average together, their characteristics were omitted from the final result entirely.

The thresholds for accuracy of the tests in this analysis should be set in the context of regulatory demands and the performance of other sensing systems - as laid out in section 2.1.1 on page 53. Requirements will thus be chosen near the thresholds of  $20 \,\mu g \, m^{-3}$  and 0.5 fractional error. The following analysis makes use of these thresholds:

- With RSE, a fit qualifies if the residual standard error between its prediction in the validation period and the reference is below 20, 25, 30, 40, or 50 μg m<sup>-3</sup>.
- For *FE*, the test is passed if the fractional error during the validation period is below 0.5, 0.75, 1 or 1.25.
- For *PCC*, the tests are passed if the Pearson correlation coefficient during the validation period exceeds a value of 0.5, 0.6, 0.7 or 0.8.

The improvements these tests had on the statistics of the instrument, coupled with the amount of coverage (that is, the fraction of calibrations that passed the test in the first place), are shown in tables 3.6 to 3.9 on page 141.

	1 Month	ı	R	SE	F	E	PC	CC	Coverage	To Spec
Gas	Test	Thres.	RSE	SD	FE	SD	PCC	SD	Percent	Percent
		20	-3.74	3.42	0.28	0.53	-0.02	0.10	22%	64%
		25	-0.39	3.50	0.17	0.34	-0.07	0.10	82%	31%
	RSE	30	-0.66	1.71	-0.02	0.15	-0.01	0.06	99%	17%
		40	-0.18	1.09	-0.01	0.06	0.00	0.03	100%	13%
		50	-0.06	1.05	-0.01	0.05	0.00	0.02	100%	11%
		0.50	2.39	2.88	0.07	0.29	-0.07	0.09	50%	0%
NO2	FF	0.75	-0.55	1.67	-0.02	0.13	-0.01	0.04	99%	20%
	FC	1.00	-0.58	1.24	-0.03	0.07	0.00	0.03	100%	18%
		1.25	-0.32	1.10	-0.02	0.05	0.00	0.03	100%	15%
		0.5	0.89	2.24	0.04	0.23	-0.03	0.08	93%	6%
	DCC	0.6	1.95	3.60	0.11	0.28	-0.06	0.11	82%	3%
	PUU	0.7	3.31	4.67	0.17	0.41	-0.03	0.17	47%	5%
		0.8	15.33	7.96	0.65	0.33	-0.16	0.26	2%	0%
Gas	Test	Level	Low is better	Low is better	Low is better	Low is better	High is better	Low is better	High is better	High is better

TABLE 3.6: Results table for the tested CAFPOLD analyses for NO<sub>2</sub> on a one month scale, showing the improvement relative to untested values, averaged across all four sensors. Each row represents a different type (and severity) of statistical test. The statistics are coloured according to how beneficial each test was to them; red implies a negative result, green a positive improvement, white means no change. With the exception of PCC, coverage and "To Spec.", lower values are better. The standard deviation of the results on this table is that of the average for the four sensors, and a higher figure implies that the test had unpredictable result between sensors (and higher values are thus coloured in darker orange). The coverage column gives the fraction of calibrations that produced results over the month, and the To Spec column shows what fraction of the time the resulting sensor could be expected to produce predictions with RSE below 20 µg m<sup>-3</sup>.

Ę	5 Months		RSE		F	E	PCC		Coverage	To Spec
Gas	Test	Thres.	RSE	SD	FE	SD	PCC	SD	Percent	Percent
		20	-3.50	3.47	0.16	0.61	0.01	0.12	22%	59%
		25	-0.87	5.00	-0.12	0.66	0.01	0.16	82%	31%
	RSE	30	-0.53	2.88	-0.04	0.27	0.01	0.07	99%	18%
		40	-0.12	1.36	-0.01	0.08	0.00	0.03	100%	15%
		50	-0.01	1.08	-0.01	0.06	0.00	0.03	100%	14%
	FE	0.50	-1.45	6.67	-0.27	0.80	0.03	0.17	50%	35%
NO2		0.75	-0.33	2.46	-0.05	0.24	0.00	0.07	99%	18%
		1.00	-0.24	1.56	-0.03	0.11	0.00	0.04	100%	16%
		1.25	-0.10	1.16	-0.02	0.07	0.00	0.03	100%	15%
		0.5	-0.33	4.94	-0.09	0.46	0.02	0.12	93%	19%
	DCC	0.6	0.10	7.67	-0.17	0.71	0.02	0.18	82%	24%
	FUU	0.7	2.78	8.73	0.14	0.66	0.02	0.21	47%	17%
		0.8	15.33	7.96	0.65	0.33	-0.16	0.26	2%	0%

TABLE 3.7: Results table for the tested CAFPOLD analyses for  $NO_2$  on a five month scale.

	1 Mont	n	R	SE	F	E	PC	C	Coverage	To Spec
Gas	Test	Thres.	RSE	SD	FE	SD	PCC	SD	Percent	Percent
		20	-1.33	6.52	0.72	0.94	0.13	0.09	33%	21%
		25	0.64	8.18	-0.04	1.13	0.08	0.13	77%	18%
	RSE	30	0.21	6.60	0.12	0.98	0.03	0.10	76%	19%
		40	1.95	7.88	0.14	0.91	0.04	0.11	94%	22%
		50	0.00	0.00	0.00	0.00	0.00	0.00	100%	10%
		0.50	-5.69	5.18	-0.93	0.79	0.13	0.09	54%	51%
O3		0.75	-1.28	7.73	-0.72	0.94	0.11	0.10	87%	35%
	FE	1.00	0.59	8.03	-0.51	0.77	0.09	0.11	81%	27%
		1.25	-0.69	8.01	-0.52	0.63	0.08	0.12	89%	35%
		0.5	-1.26	3.85	-0.36	0.42	0.06	0.06	100%	23%
	DCC	0.6	-1.78	5.22	-0.35	0.51	0.08	0.08	100%	23%
	PUL	0.7	-0.17	7.49	-0.21	0.95	0.08	0.10	99%	18%
		0.8	-0.85	10.12	-0.08	1.10	0.10	0.10	82%	38%

TABLE 3.8: Results table for the tested CAFPOLD analyses for  $O_3$  on a one month scale. The column labels are identical for table 3.7.

5 Months		RSE		F	E	PCC		Coverage	To Spec	
Gas	Test	Thres.	RSE	SD	FE	SD	PCC	SD	Percent	Percent
		20	-4.02	6.80	0.38	0.91	0.12	0.10	33%	40%
		25	-0.03	7.26	0.07	0.76	0.05	0.10	77%	18%
	RSE	30	0.34	7.08	0.01	0.62	0.02	0.09	76%	10%
		40	0.70	7.51	0.15	0.67	0.03	0.10	94%	22%
		50	0.00	0.00	0.00	0.00	0.00	0.00	100%	6%
		0.50	-4.19	8.37	-0.35	0.86	0.08	0.10	54%	42%
O3	FE	0.75	0.09	6.90	-0.17	0.79	0.03	0.13	87%	20%
	FC	1.00	1.85	7.63	0.05	0.84	0.02	0.09	81%	15%
		1.25	1.46	8.04	0.02	0.74	0.02	0.09	89%	18%
		0.5	-0.82	4.46	-0.03	0.39	0.04	0.06	100%	19%
	DCC	0.6	-1.58	6.13	-0.07	0.46	0.05	0.07	100%	23%
	FUU	0.7	-1.21	7.01	-0.01	0.65	0.04	0.10	99%	18%
		0.8	-1.95	8.97	0.03	0.78	0.07	0.09	82%	33%

TABLE 3.9: Results table for the tested CAFPOLD analyses for  $O_3$  on a five month scale.

#### 3.7.1 Discussion

The coverage statistic is an important measure of the stringency of a test, and is calculated by taking the length of the prediction arrays that passed the test threshold, and dividing it by the length of the full LTE data series. It represents the probability that a calibration will pass a particular test threshold. Choosing the best test to use is a tradeoff between between achieving a high enough coverage to be practical, and the naturally lower coverage of more stringent thresholds that generally produce better results.

The minimum acceptable coverage is determined by an experiment's practical considerations. Experiments involving multiple instruments that must be calibrated together before being deployed to another site will benefit from a calibration period that is reliably short so that all the instruments can be moved together. Experiments the use a single sensor, particularly over a short period or as part of a mobile platform (such as the roving experiment in section 6.2 on page 209), can afford longer calibration periods and coverage might not be as important. The chance p of a calibration taken over a time period  $T_c$  passing a threshold test with coverage of C during a longer validation time T is:

$$p = 1 - (1 - C)^{T_v/T_c}$$
(3.22)

Since a calibration's predictions become less useful the further away from calibration they are (see table 3.5 on page 134), it is possible that the only calibration that produces predictions better than the goodness threshold is far enough from the deployment time that its predictions are no longer superior to a calibration taken without threshold testing.

Testing data with a single performance metric tended to improve predictions with respect to that metric, but this didn't always carry over to other statistics. For example, NO<sub>2</sub> testing against an *RSE* of below  $20 \,\mu g \,m^{-3}$  improved the *RSE* markedly, but fractional error increased by 0.28, over a baseline of 1.24 on average, which is already very high, and coverage was only 22%. Testing to  $25 \,\mu g \,m^{-3}$  produces more better results, although the standard deviation of the improvements is often larger than the results themselves. The test in this case allows the specification of *RSE* lower than  $20 \,\mu g \,m^{-3} \,31\%$  of the time over both one and five months. On the other hand, the relaxed stringency means coverage is also substantially better, at 82%.

As might be expected considering the PCC of the sensors tends to remain high even months from calibration, testing based on PCC thresholds, particularly higher levels like 0.7 or 0.8, is often counter-productive for NO<sub>2</sub>. Fractional error is almost as useful as using RSE as a threshold. Like RSE, the most stringent test has low coverage and any improvement made have a high standard deviation between sensors. Testing instead to a fractional error of 0.75 was more promising, producing better slightly better statistics over one month, but almost no improvement over five months. The standard deviation of the statistical improvements is once again many times larger than the improvements themselves. The plot of this CAFPOLD analysis shown in figure 3.29 shows why this is - while some sections of the data are vastly improved (particularly for the negative months corresponding to a calibration taken at toward the end of the LTE data set), the rest of the time the improvements are marginal. For the 0.75 fractional error test, fractional error did not show the improvement that the *RSE* test did around the period of -2 to -3 months, and so the best threshold for testing data from the LTE is to an *RSE* of  $25 \,\mu g \,m^{-3}$ .

For  $O_3$ , which test to use is more clear-cut. A fractional error threshold of 0.5 produces consistently good results - for the sake of discarding a quarter of calibrations, the sensor has a better than 50% chance of performing within the specification of  $20 \,\mu g \,m^{-3}$ , and even shows improvement in PCC, in the case of data within a month of calibration the improvement is greater than the standard deviation. Figure 3.29 e) shows significant improvement in fractional error in particular over the untested baseline. d) shows that the improvement in *RSE* on average of around 4-5  $\mu g \,m^{-3}$ from tables 3.8 on page 140 and 3.9 on page 141 is quite inconsistent, explaining the large standard deviation relative to the improvement in the *RSE* statistic.

A CAFPOLD plot is shown in figure 3.29 for the best tests. While the difference is significant to the eye, the statistics tell a different story, and particularly over one month the improvement in statistics overall is quite small compared to the size of the standard deviation of those improvement and the original statistics themselves.

## 3.8 Summary

This chapter focussed on the importance of performing a pre-deployment calibration. The best calibration setup was determined from the literature and theoretically to be outside, adjacent to a reference instrument. It was determined that a week-long calibration period gives the best fit results, as shown in figure



FIGURE 3.29: CAFPOLD plot of the fit quality using the inverse fit equations and the best tests, verses time from calibration, for both  $NO_2$  (a-c) and  $O_3$ (d-f). The grey line is the baseline, which in this plot represents the average of all four untested sensors.

3.18. Bespoke calibration equations were developed for the present MOS sensors incorporated into the Zephyr prototype hardware.

One n	nonth	$RSE\mu g m^{-3}$	FE	PCC
NO2	Untested	$23.6 \pm 3.9$	$0.67\pm0.27$	$0.63 \pm 0.05$
	Test RSE $<25$	$22.9 \pm 5.3$	$0.80 \pm 0.44$	$0.57 \pm 0.13$
$\cap 2$	Untested	$23.3 \pm 3.9$	$1.11 \pm 0.65$	$0.80 \pm 0.04$
	Test $FE < 0.5$	$20.0 \pm 5.05$	$0.57 \pm 0.43$	$0.85\pm0.05$

TABLE 3.10: Table of expected performance of sensors over one month based on the Long Term Experiment dataset, obtained by averaging performance for the four individual sensors, alongside the performance achieved through the two best tests. 95% confidence intervals are given.

The best equations for predicting sensor performance used inverse terms: equations 3.18 for NO<sub>2</sub> and 3.20 for O<sub>3</sub> gave the highest performance using a simple linear regression. The statistics for the Long Term Experiment dataset using these equations are given in table 3.10. Notably the fractional error of the sensor data could be quite high (on average), and  $O_3$  predictions using these sensors were more reliable than those for NO<sub>2</sub>.

Data analysis sufficient to examine the issues of sensor degradation and the consequent deterioration of prediction quality over time were introduced. The results are shown in table 3.5 on page 134. The simple linear fit to the diminishing performance of the predictions over a five month timescale shows a degradation in *RSE* for NO<sub>2</sub> of  $2.37 \pm 0.33 \,\mu\text{g}\,\text{m}^{-3}$  and for O<sub>3</sub> of  $1.02 \pm 0.57 \,\mu\text{g}\,\text{m}^{-3}$  per month. Notably, Pearson Correlation Coefficient on average does not diminish nearly as rapidly for either gas, implying that the degradation takes the form of a linear offset, rather than a loss of sensor responsiveness.

Finally, after examining the sensor's raw performance, techniques were used to filter the calibrations most likely to produce good results. A week-long calibration was followed by a week-long validation period over which the performance statistics for that calibration's predictions were determined. Predictions that fell below a threshold of performance were discarded, necessitating longer calibration times in some cases. The procedure improves performance for predictions that occur further from calibration time, but particularly for NO<sub>2</sub> improvements were marginal or even negative for predictions within a month. For NO<sub>2</sub>, the best performance overall was gained by using a threshold RSE of 25 µg m<sup>-3</sup>. For O<sub>3</sub>, a threshold fractional error of 0.5 was best. The changes that result from this procedure are illustrated in figure 3.29.

# Chapter 4

# Understanding pollution sensors in changing environmental conditions

In the previous chapter, the usefulness of a multiple linear regression model for calibrating the MOS sensor output to a reference was demonstrated. In this chapter, more advanced analysis techniques are evaluated in order to improve the accuracy of sensor predictions.

Several techniques will be discussed in this chapter, and these can be divided into two categories:

- Alternative methods instead of simple linear regression
- Subdividing the data used for calibration based on external environmental factors and using conventional linear regression fits within those divisions.

## 4.1 Overview of advanced fitting techniques

Compensating for the sensor's non-linear response to target gases was a key theme in the previous chapter. As discussed in section 3.2.2 on page 99 and implied by the sensor datasheet[190], the simple linear regression fit with the equations used so far can only produce an approximation to the sensor's response to target gas concentration. The response of the sensor to an unknown non-target gas will likely also be non-linear. This will limit the ability of the sensor to compensate for interference in the environment when using linear equations, particularly as atmospheric gas concentrations inevitably change beyond the range experienced by the instrument during calibration.

However, trying to empirically derive a higher-order analytic model for the sensor response is unlikely to produce realistic results unless the derivation is informed by a detailed theoretical understanding[200]. Incorporating the cross-sensitivity issues exhibited by MOS devices into these models is even more complex and errorprone, given that there may be interference in the real atmosphere from multiple gases, each with their own response function, undoubtedly some of which were never considered by the sensor manufacturer. Adding additional fitting coefficients to attempt to account for the latter issue can produce a response function that either never converges onto any result at all, or is susceptible to overfitting[227].

There are machine learning techniques that go some way toward solving this problem. Of these, the most commonly used is the random forest[191][229], although an alternative that shows promise is support vector regression[191]. The advantages of these techniques is that they can not only fit an unknown non-linear function, but they can also account for cross-sensitivity. Their disadvantage is a greatly increased computation time compared to simple linear regression. A single regression fit can be completed in a fraction of a second, but a random forest of the same data can take several minutes to grow. Support vector regression reportedly takes hundreds of hours to train to a single sensor with a single dataset[191]. Further, these machine learning fits do not start from an underlying physical principle. If a sensor's response function is approximated well enough by the fitting equations, then a linear regression fit can be used to extrapolate predictions outside the range of the inputs used to produce the regression. This is not true with random forests or support vector regression, which only provide accurate results if the training dataset contains the full range of possible inputs[230].

In spite of their limitations, deriving higher-order analytic equations for fitting and using machine learning are still viable paths forward. But given the work already done on machine learning fits for air quality sensors, the subject of this chapter will instead be an alternate approach.

Instead of restricting the input variables for a calibration to the sensor voltages, humidity/temperature and the reference instrument concentrations, other variables from the environment can be used. These may be simplest to model as Boolean variables (ie. day and night). A threshold in a continuous variable like temperature can also be defined, and separate calibrations taken on either side of it. Splitting the sensor timeseries into different "bands" according to these variables allows a tried-and-tested means of fitting to incorporate additional information and partly compensate for any non-linearity. Dividing the data necessarily means that there will be fewer points to fit to during calibration, as multiple calibrations will be required for each band. The tradeoff must be borne in mind; incorporating extra variables should only be done if the improvement in prediction performance exceeds the detrimental effect of using a different calibration period, as described in section 3.4.2 on page 120.

For small pollution sensors there are several practical environmental variables along which to split the data, some of which are discussed further in this chapter:

• As shown in section 1.4.1 on page 29, nitrogen and ozone chemistry varies dramatically between day and night. While the temperature and humidity

changes of day and night Splitting the data along the diurnal cycle is perhaps the most obvious choice.

- The data can be split according to some of the variables already used for the sensor - for example, making separate fits for high and low humidity and merging them together. This stitching together of different linear fits is a relatively crude way of approximating non-linear functions, but it has the advantage of being relatively fast.
- Wind direction is another way of dividing the data. Different processes emit different mixes of pollutant gases, and the motion of the air in general determines whether those pollutants from that source reach the sensor, how much deposition occurs and any chemical mixing or reactions that occur on the way. This is a complex process and modelling of the wind environment along with the physical geometry of the sensor surroundings are needed to fully account for it[231]. But any correlation of pollutants by wind direction could still be used to section data[232], and potentially improve sensor fitting.
- The calibration period itself can be split into two, making separate calibration periods on either end of the experiment or deployment. If the sensor degrades or some environmental change throws off the calibration at the start of the experiment, combining that fit with the one at the other end of the experiment might compensate for those changes and improve prediction performance, essentially by accounting for the transition between the two different sets of calibration coefficients.

Using machine learning techniques to calibrate small pollution sensors is an active area of research with the potential to improve performance significantly. However, subdividing the data to account for environment variables has not so far received a thorough examination in the literature regarding these sensors, and so it will be the focus for this chapter. However, for the sake of comparison, a discussion of a
random forest machine learning techniques in the terms used in this thesis is made in appendix C.

# 4.2 Correcting for day/night chemistry

Section 1.4.1 on page 29 described some of the significant differences between day and nighttime atmospheric chemistry. Photons are a reagent of sorts and their sustained absence during the night transforms many chemical equilibria. The Leighton relation does not always hold in practice during the day, for example in the immediate vicinity of strong pollution sources, but at night the lack of UV photons makes it irrelevant and instead the nitrogen in the atmosphere can take other forms which may interact differently with MOS sensors.

In addition to this, changes in temperature at night, stratification of the troposphere in the absence of convective forces [233], and changes in the concentration of other atmospheric compounds in the absence of UV light [226] are all factors that can radically alter the composition of the atmosphere, and thus might have an effect on sensor performance outside of the concentration of NO<sub>2</sub> and O<sub>3</sub> that our models thus far have been fitting to. In this section the magnitude of such effects on the sensor's performance will be investigated.

#### 4.2.1 Method

The simplest way of determining the intensity of this effect is to directly compare fits of sensor data to the AURN, taken during the day and night. The AURN station records Photosynthetically Active Radiation (PAR) in  $W/m^2$ , a reading proportional to the total intensity of incident sunlight. Along the axis given by this metric it is possible to split data accordingly - an arbitrary boundary of 40 W m<sup>-2</sup> is used throughout this section to form a clear delineation, as during the day at the AURN in Leicester PAR regularly exceeds  $300 \,\mathrm{W m^{-2}}$ . A PAR below this level for the purpose of this analysis indicates twilight and the start of night time.

As discussed above, splitting the data in two will require twice the calibration data to achieve comparable fit quality - as demonstrated in section 3.4.2 on page 120, doubling the standard 7-day fitting window does not significantly improve fits, but halving it can increase error substantially. This is complicated by the fact that a longer calibration period will be able to fit around gaps in the data from the long term experiment far less easily. Consequently the fitting period for all these techniques will be two weeks long, and they will be compared to a baseline 2-week fitting period instead of the results achieved in the previous section. The only stretch of data suited to this experiment was taken in June, where in the UK the day was naturally longer than the night.

Five CAFPOLD analyses were performed using various types of sectioning for the two-week-long calibration period. These were:

- *Baseline*: Fits are produced from the entire calibration period, as normal. To keep the baseline fit at a comparable quality to the other ones in this experiment, only every other data point is used for calibration (giving the input data a period of two hours).
- Day only: Fits are produced from only those calibration periods with PAR greater than 40  $W/m^2$ . They are then applied to predict concentrations for the whole period.
- Night only: Fits are produced when PAR is less than 40  $W/m^2$  and used for all predictions, day or night.
- *Diurnal*: Two fits are produced during the calibration period, one each for day and night. These then predict data that was taken during the day and night respectively.

• *Reversed*: Fits are produced during both day and nighttime as above, but used to predict data during night and day respectively. Intuitively this should be the worst performing sectioning method, if day verses night has a systematic effect on sensor response.

After being divided in the first stage of the analysis into fourteen-day sections (as described in section 3.4.3 on page 124), the data sections were discarded if each section individually contained less than six days worth of measurements - combined, each would have at least twelve days of data. Comparing the consequences of splitting the data in this way will reveal the extent to which differences in day and night atmospheric chemistry affect the MOS sensors.

Attaining a minimum 12-day long calibration period to register a fit with the CAF-POLD analysis was difficult in the LTE dataset, which has several discontinuous sections. The CAFPOLD analysis with a weeklong calibration produced results over  $\pm$  5 months, but for the two week period no results were returned beyond a month and a half from calibration. With a more reliable experimental setup, longer datasets using this analysis could be obtained. Nonetheless the data series is long enough to compare performance with the baseline analysis (which did not discriminate between day and night).

#### 4.2.2 Results and discussion

In the following analysis and comparison for NO<sub>2</sub> only data from sensors a and b are used. Figure 4.1 shows the result from the CAFPOLD analysis using the diurnal sectioning, demonstrating this behaviour, and these two sensors perform similarly for the other sectioning methods. The NO<sub>2</sub> results from sensors c and d are similar to those of a and b for those predictions one month earlier than the calibration time (see the patch at -1 month in figure 4.1), but are extremely poor for predictions one month beyond the calibration point, with *RSE* nearly  $100 \,\mu g \,\mathrm{m}^{-3}$  higher for sensor d shown in panel a) - this spike is over four times the acceptable margin of error as given in section 2.1.1 on page 53. The poor performance of the OXc and OXd sensors in certain circumstances is a recurring theme - for further discussion see section 3.6 on page 133 regarding figure 3.28 a), where sensor d has especially poor performance in *RSE* around the -4 month mark.

To clearly distinguish different sectioning methods, the predictions from all used sensors (a and b for  $NO_2$ , a b c and d for  $O_3$ ) are averaged together to produce a single CAFPOLD timeseries. The results of this procedure for the different sectioning methods is shown in figure 4.2. Visually the difference between them is marginal on all statistics, less than the standard deviation of the CAFPOLD timeseries. Figure 4.3 shows more detail by plotting the difference between the different sectioning methods and the baseline two-week calibration, with the baseline set at zero. This procedure removes the variation over time inherent to the CAFPOLD analysis due to any degradation in the quality of the sensor predictions.



FIGURE 4.1: Example of a CAFPOLD analysis using two weeks of calibration data with the diurnal sectioning. The output for each sensor is shown as a separate trace. Note that both axes of the graphs are different to that of previous CAFPOLD analysis diagrams, spanning one and a half months of time in the x axis and different scales for the RSE and FE in graphs a) b) and d) e). Sensors c and d exhibit poor performance, particularly around +1 month in the RSE graph a). The blue vertical lines are the period of calibration.

Figure 4.2 shows that the performance of the fits is very similar over the different sectioning methods, implying a null result. It is also comparable to that of the one week long calibration. The mean RSE for the "diurnal" sectioning, produced by averaging over the timeseries, is  $19.09 \pm 8.56 \,\mu\text{g}\,\text{m}^{-3}$  for NO<sub>2</sub>, and for O<sub>3</sub> is  $22.71 \pm 8.24 \,\mu\text{g}\,\text{m}^{-3}$ . While this is respectable, it is not directly comparable to the figures for the CAFPOLD analysis using a one week calibration shown in table 3.10 on page 144, especially due to the high uncertainty in these statistics (which is obtained from the standard deviation of the time average). The mean fractional error is  $0.428 \pm 0.280 \,\mu\text{g}\,\text{m}^{-3}$  for NO<sub>2</sub> and  $0.364 \pm 0.106 \,\mu\text{g}\,\text{m}^{-3}$  for O<sub>3</sub>. *PCC* for the NO<sub>2</sub> timeseries is quite poor at  $0.485 \pm 0.278$  on average, compared to a normal CAFPOLD *PCC* for NO<sub>2</sub> of between 0.6 and 0.65. For O<sub>3</sub> it is better at  $0.825 \pm 0.051$ . From table 4.1 on page 157, the "diurnal" method gives very



FIGURE 4.2: Graph of CAFPOLD results for  $NO_2$  and  $O_3$ . The graph uses the averaged predictions of the sensors used in the experiment. Each colour of line corresponds to a different sectioning method. Once again the vertical scales here are different to the other CAFPOLD analyses in this work. The horizontal black line indicates  $20 \ \mu g \ m^{-3}$  target performance.

similar results to the others overall, with almost all differences being smaller than the 95% confidence interval. This similarity implies that the poor PCC for NO<sub>2</sub> is a consequence of the longer calibration period (as shown in section 3.4.2 on page 120), rather than poor correlation caused by the splicing together of data with different fits applied to it.

The similar performance of the sectioning methods is more clearly shown in figure 4.3, and the statistics of these data in table 4.1 on page 157. The "diurnal" sectioning method unsurprisingly gives slightly better performance than the "reversed" sectioning method for both target gases. In the case of NO<sub>2</sub> RSE, the "reversed" sectioning produced predictions  $1.71 \pm 2.65 \,\mu g \,m^{-3}$  worse than the diurnal method, which is a little over one standard deviation of difference. This is intuitive, but not a strong enough effect to be statistically significant, and is not consistent over the other performance metrics. More crucially, no sectioning method was significantly better than the baseline, which assumed no distinction in prediction between night and day. For  $O_3$ , the night only sectioning is the only one that on average improved over the baseline, albeit by less than the confidence interval. To emphasise this point on how marginal the reduction in *RSE* is numerically, note that the "night only" sectioning was occasionally capable of outperforming the "diurnal" sectioning (which might be expected to be optimal) for both gases - see figure 4.3 a) and d), particularly for  $O_3$  at -0.5 to -1.5. In panel a) at +1.0 to +1.5 months all sectioning methods except for diurnal show *RSE* well above the baseline (which in the figure is the flat grey line set at zero), and diurnal sectioning vacillates above and below the baseline level.



FIGURE 4.3: CAFPOLD graph of difference between sectioning methods and the baseline 14-day calibration. Each colour of line corresponds to the predictions using a particular sectioning method minus the baseline prediction. The scale is zoomed in to make the difference in methods clearer.

		$ m RSE~\mu gm^{-3}$	FE	PCC
$NO_2$	Day only	$+0.90 \pm 1.41$	$+0.03 \pm 0.08$	$-0.09 \pm 0.11$
	Night only	$+0.40 \pm 2.39$	$+0.01 \pm 0.07$	$0.00\pm0.06$
	Reversed	$+1.63 \pm 2.53$	$0.05\pm0.07$	$-0.11 \pm 0.09$
	Diurnal	$-0.08 \pm 0.80$	$-0.01 \pm 0.03$	$+0.01 \pm 0.04$
O <sub>3</sub>	Day only	$+0.59 \pm 1.17$	$+0.03 \pm 0.05$	$-0.02 \pm 0.03$
	Night only	$-0.12 \pm 1.04$	$-0.01 \pm 0.03$	$+0.01 \pm 0.02$
	Reversed	$+0.81 \pm 0.69$	$0.02\pm0.03$	$-0.02 \pm 0.01$
	Diurnal	$+0.04 \pm 0.62$	$0.00\pm0.01$	$+0.00 \pm 0.01$

TABLE 4.1: Table of statistics from the CAFPOLD analyses of different sectioning techniques, giving the relative difference between the sectioning technique and the baseline. The statistic is obtained by averaging this difference over the whole timeseries. The error incorporates the uncertainty due to averaging between sensors and over time. For all metrics, lower is better except for PCC, where the best results are those closest to 1.

#### 4.2.3 Summary of outcomes

The sectioning methods clearly do have a modest impact on prediction performance. But this impact is so small that no sectioning method performed consistently better than the baseline prediction in *any* metric by an amount greater than the 95% confidence interval. Thus the use of this technique to improve fit predictions cannot be recommended, especially given the practical considerations involved in doubling the length of the calibration period. However, this negative result proves that the difference in atmospheric chemistry between night and day does not introduce any statistically significant complications when using MOS sensors to measure NO<sub>2</sub> or O<sub>3</sub>.

### 4.3 Transitional fits

As shown in section 3.6 on page 133, sensor fits can vary wildly in quality, with predictions sometimes becoming markedly less accurate further away from the time of calibration. The purpose of a transitional fit is to adjust a sensor's calibration values to incorporate two calibrations, performed at the beginning and the end of a calibration period. During an instrument deployment, rather than making the choice to use a calibration either at the start or the end of the experiment, both calibrations can be interpolated, showing more influence for one calibration or the other depending on the time at which a prediction must be made, and potentially improving accuracy.

This is not a new technique - Smooth Transition Auto-Regressive (STAR)[234] modelling is often used in economics to transition between two time domains governed by different statistics. For this work the transition will be a simple linear interpolation, rather than the exponential formulae used by STAR. This different use of interpolation could be significant, and whether the linear transformation is a more realistic way of compensating for sensor interference due to changing gas environments, or the degradation of the sensors over time, is a topic for future work.

Equation 3.5 in section 3.2.1 on page 97 showed the response (in terms of changing resistance) of a Metal Oxide gas sensor to the environment. A more general version of this equation is:

$$y = f(x_1, x_2, x_3 \dots x_i) \tag{4.1}$$

where y is the desired output variable (in our case, gas concentration) and f is a function of i input variables x. A simple linear fit uses a function with this form:

$$f = \sum_{i} k_i x_i \tag{4.2}$$

where the k terms are, of course, the calibration constants found during the fitting process.

When dealing with a point of raw sensor data  $x_t$  taken at time t, between the time at which the first calibration ended  $t_s$  and the second began  $t_e$ , a linear interpolation can be made between the two calibrations by transitioning the coefficients of each ( $k_s$  and  $k_e$ ), taking the example from equation 4.2 in the context of equation 4.1:

$$y_t = \sum_{i} \left( k_{si} \left( 1 - \frac{t - t_s}{t_e - t_s} \right) + k_{se} \left( \frac{t - t_s}{t_e - t_s} \right) \right) x_i \tag{4.3}$$

The linear transition in this equation is accomplished by replacing the coefficients  $k_i$  with a simple function  $k_i(t)$ :

$$k_i(t) = k_{is} \left( 1 - \frac{t - t_s}{t_e - t_s} \right) + k_{ie} \left( \frac{t - t_s}{t_e - t_s} \right)$$

$$(4.4)$$

Using the inverse multiplicative equations from section 3.4.1 on page 117 that proved most useful for fitting to NO<sub>2</sub> and O<sub>3</sub> gives

$$[NO_{2}] = k_{c}(t) + \frac{k_{1}(t)}{V_{OX}} + \frac{k_{2}(t)}{V_{OX}}RH + \frac{k_{3}(t)}{V_{OX}}T + \frac{k_{4}(t)}{V_{OX}} + \frac{k_{5}(t)}{V_{OX}} + \frac{k_{5}(t)}{V_{OX$$

for  $NO_2$ , and

$$[O_3] = k_c(t) + \frac{k_1(t)}{V_{O3}} + \frac{k_2(t)}{V_{O3}}RH + \frac{k_3(t)}{V_{O3}}T$$
(4.6)

for  $O_3$ .

#### 4.3.1 Method

To determine the improvements possible using transitional fits, a section of long term experiment data was analysed in two different ways:

- To test the transitional fit, two separate calibrations fits were taken, over week-long periods four weeks apart. A transitional fit prediction was made during the intervening time.
- The control was to take two fits over the separate one week-long periods, as above, and then produce four week long predictions for each.

In both cases, the quality of the prediction is measured over the whole span of the prediction - the "validation period", which in both cases is four weeks long. The timestamp for this analysis reads from the start point of the first calibration period.

The question to be answered is this: does using a transitional fit improve the sensor's performance compared to using a single fit from the start or end of the validation period? How consistently does the transitional fit perform? In practice, whenever a sensor is being used on a mobile deployment, it will be impossible to check the accuracy of the sensor during the period it is away from the reference instruments at the AURN station. Consequently, if a decision is made to use transitional fits, the consistency by which they improve results is just as important as the average improvement over using a single fit.

#### 4.3.2 Results

The transitional fit was performed over a large section of LTE data, and in figures 4.5 and 4.6 the quality achieved over time is plotted. This is not a CAFPOLD analysis as there is no averaging over different sets of calibrations to produce a quality metric output. Instead the graphs should be evaluated according to figure 4.4.

Figures 4.5 and 4.6 show the results of this analysis, with the calibration from both before and after the evaluation period (green and purple respectively), as well as the transitional fit in black. Table 4.2 on the following page shows performance numbers for each of the individual sensors, including those omitted from 4.5 because of poor performance.



FIGURE 4.4: Timing diagram for the transitional fit plots in this chapter. The time represents the beginning of the block of starting calibration, evaluation, and ending calibration. Each point on the output quality plot shows the quality achieved during the evaluation period for the respective time.

	Start					End				Transitional								
	RSE		FE		PCC		RSE		FE		PCC		RSE		FE		PCC	
Sensor	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
NO2a	27.7	7.7	0.74	0.12	0.33	0.08	18.1	1.7	0.43	0.09	0.44	0.06	20.0	3.1	0.49	0.03	0.44	0.05
NO2b	28.2	8.8	0.78	0.13	0.38	0.10	23.9	5.9	0.42	0.07	0.34	0.12	22.0	6.2	0.49	0.03	0.46	0.11
NO2c	113	71.8	1.05	0.32	0.05	0.09	111	82.9	0.89	0.46	0.04	0.10	97	51.4	0.84	0.26	0.08	0.13
NO2d	163	122.1	1.63	0.82	0.19	0.06	79	64.1	0.82	0.49	0.22	0.10	124	84.1	1.16	0.49	0.21	0.08
O3a	32.1	4.4	0.41	0.10	0.78	0.04	32.0	5.5	0.48	0.24	0.71	0.15	30.7	1.5	0.41	0.11	0.77	0.05
O3b	28.4	4.2	0.36	0.09	0.81	0.04	32.3	8.0	0.50	0.25	0.67	0.22	28.3	2.5	0.39	0.12	0.78	0.08
O3c	25.8	4.7	0.37	0.06	0.84	0.03	29.8	5.6	0.50	0.23	0.75	0.16	26.4	2.1	0.39	0.10	0.82	0.06
O3d	28.6	4.4	0.45	0.08	0.80	0.03	32.5	6.8	0.53	0.26	0.71	0.17	28.4	2.4	0.43	0.12	0.79	0.07

TABLE 4.2: Table of performance statistics for each of the sensors using the different fitting methodologies. The statistics are calculated over the whole dataset produced by the above analysis, shown in figures 4.5 and 4.6. The statistics in the transitional fit section are coloured according to whether the transitional fit beat both other fit methods (blue), performed better than the average between the two (green), were worse than average (yellow) or were the worst option (red). The results from the substandard OXc and OXd sensors are highlighted in red in the leftmost column.

#### 4.3.3 Discussion

In the NO<sub>2</sub> graphs of figure 4.5, the fractional error and RSE for both sensors in question was quite bad for calibrations taken during the first two weeks. One of the sensors gave predictions from the end of the validation period that was sometimes better than the 20 µg m<sup>-3</sup> specification when used for the same data. In the first two weeks the transitional fit was on average 3-4 µg m<sup>-3</sup> worse than the



Beginning of validation period

FIGURE 4.5: Evolution with time of residual standard error, fractional error and Pearson correlation coefficient for the three fitting methodologies, for prediction  $NO_2$  concentration. The two lines in each colour are for sensor a and b; sensors c and d performed very poorly and are not shown here. The dates are all in 2016.

end-period calibration, but far and away better than the start calibration, which often exceeded an RSE of  $40 \,\mu g \,m^{-3}$ . The fractional error graph tells the same story even more clearly. Toward the end of the timeseries from the 19th of June, the end calibration for one of the sensors starts producing worse predictions, and again, the transitional fit is not affected.



Beginning of validation period

FIGURE 4.6: Evolution with time of residual standard error, fractional error and Pearson correlation coefficient for the three fitting methodologies, for prediction  $O_3$  concentration, with multiple traces for sensors a, b, c and d.

For both calibrations, however, the PCC was very low for every fit - frequently less than 0.5, and far weaker than the week-long evaluation period results shown in the previous chapter (see table 3.10 on page 144).

The  $O_3$  graph in figure 4.6 tells a similar story; where either the start or end period fits produces poorer statistics (for example the end fit around the 13th of June), the transitional fit is not affected nearly as badly. Instead of merely sitting between the start and end fit in terms of quality, the transitional fit gives consistently better results than the average of the start and end fits - although it does not always perform better than the *best* of the start or end fits.

Table 4.2 on page 161 shows that remarkably, for every metric, the transitional fit has statistics that are better than the average of either the start or end fits, except in the case of the sub-standard c and d sensors fitting NO<sub>2</sub>. The transitional fit had the better (that is, lower) *RSE* than the start or end fits in the case of NO<sub>2</sub> from sensor b, and all the O<sub>3</sub> predictions save for sensor c - that's four of the six fully-performing sets of sensors.

Interference from non-target gases or changes in the weather might cause a calibration to be less relevant to the rest of an experiment's dataset, indeed, figure 4.6's dip in performance for the end calibration around June 13th shows particularly clearly that all four sensors suffered in performance at the same time and in roughly the same way. But the transitional fit seems able to compensate for this phenomenon better than merely averaging the results of either calibration method.

#### 4.3.4 Summary of outcomes

The takeaway is that in an experiment that doesn't have regular access to a reference instrument to compare to, the transitional fit would provide a more accurate way of determining gas concentration. For quantification, see table 4.3 on the next page which shows the relative difference in performance statistics.

It is important to consider that in a deployment in the field, there is no check against a reference dataset to determine the best fit, and good performance will depend on a field environment having similar characteristics to the calibration environment. In table 4.3 on the following page therefore, the realistic comparison is with the average quality of the start and end period fits. With this comparison, the transitional fit gives consistently better performance, in almost every case by a

		$ m RSE~\mu gm^{-3}$	$\rm FE$	PCC
	Worst case	$+6.92 \pm 6.20$	$+0.27 \pm 0.18$	$-0.10 \pm 0.09$
$NO_2$	Best case	$-0.05 \pm 4.60$	$-0.07 \pm 0.09$	$0.06\pm0.07$
	Average	$+3.44 \pm 3.00$	$+0.10 \pm 0.05$	$-0.08 \pm 0.09$
	Worst case	$+3.19 \pm 5.01$	$+0.10 \pm 0.05$	$+0.08 \pm 0.07$
$O_3$	Best case	$-0.26 \pm 3.36$	$+0.01 \pm 0.08$	$-0.02 \pm 0.10$
	Average	$+1.72 \pm 1.85$	$+0.05 \pm 0.03$	$+0.03 \pm 0.03$

TABLE 4.3: Relative increase in performance between transitional fit and either the worst performing of the start and end period calibration for the sensor in question, the best of them, or the average of the performance metrics of the two calibration periods. A positive figure indicates that the transitional fit improved performance. 95% uncertainties are computed from the standard deviation of the difference between the two dataseries. The figures here are averaged between the individual sensor results, although the NO<sub>2</sub> c and d sensors were not used in this average, partially accounting for the higher uncertainty of NO<sub>2</sub> statistics.

margin greater than the 95% confidence interval - for the  $O_3 RSE$  and  $NO_2 PCC$ , the improvement is positive and similar in size to the 95% confidence.

While in most metrics the transitional fit is slightly worse than the performance of the best of the individual sensor calibrations, the deficiency is in no case larger than the 95% confidence interval, and in the case of NO<sub>2</sub> *PCC* it beats it.

Some of the regions of poorer fit performance have a substantial period compared to the length of this analysis, as can be seen from the first eight days of RSE in figure 4.5 particularly. As the LTE dataset did not return a large section of data suited to testing transitional fits in this way a longer timeseries analysis that would fully capture the nature of this variation is not possible, and a longer experimental period would improve the confidence of conclusions regarding the transitional fit.

Except in cases where one of the constituent calibrations produces predictions that are well beyond the margins for error for either the reference (or other fits for that matter), this new analysis method is capable of producing better results than constraining the measurements to just a single calibration at either the start or end of the field period.

## 4.4 Wind direction and calibration

The airborne chemical environment around any instrument is a combination of gases that have been transported from regional sources, and those that have been emitted nearby and travelled almost directly to the instrument. Local source attribution is an important question for small sensors into which work has already been done[90]. In the atmosphere, uneven distribution of point sources of pollutant gas will directly affect the instruments intending to measure them by causing a correlation between wind direction and sensor performance during calibration[232].

The data used in this analysis is from the five month period between June and November 2016.

The objective is to determine whether there are any clear sources of nearby pollution that would cause a significant difference between the measurements of the AURN instruments and the MOS sensors, and potentially poorer MOS sensor predictions as a result - taking into account any sampling error from the immediate environment of the sensor. If an unknown interfering gas is being emitted from a source nearby, excluding data points representing air from a calibration could improve performance.

#### 4.4.1 Experimental conditions

The University of Leicester DEFRA air pollution monitoring station is assigned a simulated wind speed and direction derived from measurement instruments across the city of Leicester[148]. The wind data has 60 second intervals, in contrast with the ten second measurement interval of the long term experiment, and the hourly measurements of chemical composition produced by the AURN station.

#### 4.4.2 Analysis

To work with the long term experiment, then, the wind speed and wind direction vectors must be averaged to the hour. This is complicated by the fact that the AURN reports wind vectors using radial co-ordinates (wind speed and direction), and these need to be converted to Cartesian co-ordinates for this analysis. Figure 4.7 shows a comparison of the two ways of doing this: converting before averaging the resulting Cartesian wind vector, and averaging over the radial co-ordinates before converting them. The undesirable artifacts produced by averaging the radial co-ordinates is clear in this view. Distinct plumes are visible in figure 4.7, reflecting that the wind is more likely to blow over the AURN station from some directions than others.

To compare the quality of fit predictions with respect to the wind direction, a baseline fit where the wind speed was low must be made. To do this, data for which the wind speed was within the lowest quintile was used - for the LTE dataset this was less than  $0.242 \,\mathrm{m\,s^{-1}}$ . This data represents a "central region" on a wind rose diagram, those data points closest to completely motionless air. Other studies have considered wind speeds of less than  $2 \,\mathrm{m\,s^{-1}}$  to preclude regional transportation[90].

A normal linear regression fit to both NO<sub>2</sub> and O<sub>3</sub> was then taken using the inverse equations 3.18 and 3.20, using this central patch of data (which spans in time throughout the entire dataset). This fit was used to predict data values for the entire dataset. The difference between the resulting prediction and the reference values, for a sensor measuring NO<sub>2</sub> and O<sub>3</sub>, is plotted in figures 4.8(a) and 4.8(b) as a function of the windvector.



FIGURE 4.7: Wind vector data, using the two averaging methods. Wind direction is given in terms of source, so a northerly wind blows air from north to south. The positive Y axis of the graph points north.

#### 4.4.3 Results

This plot is striking because of the clear pattern of over and underestimation depending on the wind direction. Within the wind rose, spurs of blue or red coloured points indicate wind directions that consistently correlate to over or underestimation of the target gas. Particularly there is underestimation of  $NO_2$  in a broad angle of wind coming from the north and north-west. The effects of wind on  $O_3$  seem more moderate. Graphs of this kind are consistent with the hypothesis that interfering gases having specific, localized sources that can be carried to the AURN station through the wind.



FIGURE 4.8: Scatter plot of wind vectors, coloured according to the difference between predicted and actual concentrations of a)  $NO_2$  and b)  $O_3$ . Sensor a from both cases is the one shown. Anomaly refers to the predicted gas concentration minus the reference concentration. The segments drawn on the diagram each contain the same number of points, and indicate density of wind vector points that might not be apparent from the points alone.

The violin plots of figures 4.9 and 4.10 quantify the effect that was clear to the eye in figures 4.8(a) and 4.8(b). The plumes in the wind rose, which represent narrow bands of direction but seem to be less dependant on speed. For this reason in figures 4.8(a) and 4.8(b) the data is divided with respect to wind direction into eighteen 20-degree slices, excluding the low wind speed patch used for calibration.

For  $NO_2$ , the tendency for the predictions to underestimate  $NO_2$  is represented strongly in the violins for angles between zero and ninety degrees (north-northwest, as in the earlier scatter plot 4.9). The largest sector of data between 0 and -20 degrees, with 65 data points for each sensor, has a mean underestimation of  $13.1 \,\mu g \,\mathrm{m}^{-3}$ , albeit with a standard deviation of  $30 \,\mu g \,\mathrm{m}^{-3}$ . The largest underestimation was for -20 to -40 degrees,  $27.3 \,\mu g \,\mathrm{m}^{-3}$  with standard deviation of  $28.3 \,\mu g \,\mathrm{m}^{-3}$ . This offset was consistent for both sensors.

For  $O_3$ , the largest sector with 65 data points has an anomalous underestimation of  $10.8 \,\mu g \,m^{-3}$ , but with a standard deviation of  $11.7 \,\mu g \,m^{-3}$ .

The largest issue in drawing firm conclusions about the connection between wind speed and anomalous concentrations is the high uncertainty accompanying each data point. While visually patterns can be easy to distinguish particularly for NO<sub>2</sub>, statistically there is enough variance that the 95% confidence interval is roughly twice the size of the largest anomalous underestimation for both target gases. But within each segment there is some consistency between the paired sensors. The standard deviation of their difference is  $3.13 \pm 10.34 \,\mu g \,m^{-3}$  for NO<sub>2</sub> and  $2.19 \pm 4.44 \,\mu g \,m^{-3}$  for O<sub>3</sub>, far less than would be expected if the variance was solely down to random error in the sensors, as opposed to being mostly due to the environment.

#### 4.4.4 Further investigation

There is still a suggestion that compensating for sources of interference that can be distinguished through wind direction. In accordance with the sectioning techniques used in this chapter, the data is divided into six segments with respect to wind direction, each with an equal number of data points within them (as opposed to the 20° intervals shown in figure 4.9). The fit from the lowest quintile of wind speed is used to make predictions for the dataset. The *RSE* between prediction and reference instrument is shown in figure 4.11.



FIGURE 4.9: Violin plots of difference between predicted and actual data, divided according to the prevailing wind direction. For each violin plot, the thicker the band between the vertical axis and the edge of the violin, the more data points lie within the region quantified by the vertical axis. There are multiple violins, each one representing a 20 degree segment of the wind rose. Some violin strips are very narrow and represent very few overall data points in this wind direction, the lowest having five. The (green) left and (purple) right sides of each "violin" represent the two different sensors, a and b. The darker coloured beads indicate the mean of the anomaly for that sensor for wind direction within the respective angle.



FIGURE 4.10: As above, but for  $O_3$ .



FIGURE 4.11: A wind scatter plot for  $NO_2$  sensor a, overlain with boundaries chosen such that the data is divided into six segments, excluding the innermost quintile of wind speed data, and each containing the same number of points. The *RSE* value of each segment is superimposed. The colour of the segments emphasises this (ranging through green for the lowest *RSE* to red for the highest one).

However, dividing the data up in this way reveals another problem with trying to compensate for changing wind direction. For any calibration, there needs to be a period of time where, however the data are divided, each division contains enough data points for a reliable fit to be produced. For the inverse  $NO_2$  equation, this hard minimum is six (because of the six unknowns), and for  $O_3$  it is four - compare this to the typical calibration data in CAFPOLD analyses, which is at least six day's worth - that's 144 points of hourly data.

Wind does not work this way, however. Wind can blow consistently in one direction

for hours, and if the wind is filling one of the wind segments with data points, it isn't doing so for any of the others. Figure 4.12 shows this effect clearly - each trace is either for the core (in black) or one of the segments depicted above (in colour). To create a fit in the separate segments to use for predictions of unknown data, all of these traces must be above the critical level of six data points, preferably much higher. But this is not the case for almost all of the data series. This issue makes the methodology outlined in this chapter unsuitable for correcting sensor measurements for bias due to wind direction.

The use of simulated wind data is an obvious point for future improvement. Both the local geometry of the buildings and the natural turbulence in the atmosphere make in-situ wind sensors more desirable for analyses of this kind.



FIGURE 4.12: Timeseries showing number of data points in each sector for a week-long section of data, throughout the timeline of the LTE in 2016. Marked horizontally in grey is the critical level below which, for the inverse  $NO_2$  equation, no fit can be made.

#### 4.4.5 Summary

The dependence for sources of interfering gases on wind direction has been described in this section, but any strategies stemming from this insight are hard to implement for practically our calibration procedures.

Indeed, as a consequence of the way the wind moves, dividing up the data by wind direction leads to segments that contain too few points to produce a good fit. Excluding data when the wind direction indicates that it is likely to produce poor fits is a possibility, but it may be less trouble (and yield better results) to use the goodness tests described in section 3.7 on page 137.

# 4.5 Timeshift error analysis

The importance of accurate time on the sensor's clock during calibration is obvious. The extent of the degradation in fit quality caused by timing errors are interesting by themselves and might be useful in future experiment design. Additionally, this method may be used to discover timeshifts or inaccurately reported timestamps as a result of mistakes (for example, failure to correct between UTC and BST).

#### 4.5.1 Method

This analysis used a section of the Long Term Experiment data between June 1st, 2016 and November 15th, 2016. For this set of data, the following steps were repeated:

• A time shift was added to the Long Term Experiment timestamps, moving them in ten second increments through 24 hours, both into the future and the past. This is the same period as the measurement interval for the Long Term Experiment sensor.

- The DEFRA data is reported hourly. The LTE data was divided into hourlong sections starting at each DEFRA data timestamp, and averaged.
- A series of calibrations were made between the shifted LTE data and the DEFRA reference instrument, offset by ten seconds each time. As usual, for each point in time a weeklong fit was taken, and data was then predicted during the subsequent week. The goodness metrics of this data were calculated for each fit Residual standard error, Fractional error and Pearson Correlation Coefficient.
- The goodness metrics for these fits were averaged together over the whole timeseries, and the resulting values represent the average goodness possible for calibrations when the timeseries is shifted by a set amount.

The results of this analysis are shown in figure 4.13



Averaged effects of timestamp errors

FIGURE 4.13: Results of the timeshift analysis. The x data is in hours, with zero representing zero timeshift.

#### 4.5.2 Discussion

As might be expected, around a timeshift of zero produces the best results. The quality metrics do not, however, follow a simple normal distribution. Rather, the bell of a normal distribution lies at zero timeshift, but becomes indistinct beyond a shift of about four hours. The *PCC* falls until it hovers around zero beyond about six hours of time shifting, but it has a prominent peak that makes it probably the easiest metric to use for diagnosis. This timeshift analysis could be used to identify clock errors, where two datasets that have a significant timestamp offset could be reconciled.

# 4.6 Summary

Four different analysis were performed in this chapter. Three of these explored ways of improving fits, and one is more a useful diagnostic tool for ensuring that any clock errors can be caught.

Of these, the transitional fit is a promising way to improve the accuracy of predictions for field experiments. For roving experiments particularly it proved itself capable of getting better results than the average of taking single fits during calibration periods at the start and end of the experiment, as shown in table 4.3 on page 165. Compared to the average of calibrating once at either the start or the end of a four week test period, the mean improvement in *RSE* for NO<sub>2</sub> was around  $3.14 \pm 3.00 \,\mu\text{g}\,\text{m}^{-3}$ , and for O<sub>3</sub> were  $1.72 \pm 1.85 \,\mu\text{g}\,\text{m}^{-3}$ .

The analysis using wind was disappointing due to the difficulties in implementing practical improvements using sectioning techniques, although the existence of a non-uniform wind distribution, and non-uniform anomalies between prediction and reference with it, hints at the possibility of using different analysis methods to compensate for this issue. The day/night analysis also yielded a marginal improvement compared to the standard methods introduced in chapter 3. Sectioning the data into day and night, using individual fits for each and making predictions using ambient light as a variable came at the cost of doubling the calibration time needed. The level of

improvement for this technique is indistinguishable (see table 4.1 on page 157): for NO<sub>2</sub> a marginal  $0.08 \pm 0.80 \,\mu g \,m^{-3}$  reduction in RSE, and for O<sub>3</sub> an equally slight  $0.04 \pm 0.62 \,\mu g \,m^{-3}$  increase. This suggests that the change in atmospheric chemistry between day and night does not significantly affect MOS sensors. These are the best conclusions that can be drawn from the LTE data set, where a longer and more uninterrupted data set would have produced clearer results.

# Chapter 5

# Characteristics of MOS sensors and their implications for instrument design

As was described in section 2.7 on page 81, the hardware design of an instrument has a significant effect on the fidelity of the sensors inside. While the more obvious design issues mentioned in that section were avoided in the prototypes used in this work, there are some design elements that are not universally applied in air quality instruments, and these should be investigated. Specifically:

- Whether to use fans to ventilate the sensor housing.
- The "warm-up" period over which a MOS sensor becomes stable after the heating element is turned on.
- The extent that variations in the manufacturing process affect the response of different sensors of the same type.

This chapter describes the experiments used to find answers to these questions. Section 5.1 on the next page examines the effects of using a fan to actively draw air into the sensor casing, as opposed to mounting sensor elements externally. Section 5.2 on page 183 examines the time it takes for a sensor to become useful after being powered up, including a model of the active sensor surface area. Section 5.3 on page 190 examines MOS sensor variation, both in terms of raw output and correlation between adjacent sensors. This is possible on a large scale due to the instrument containing multiple sensor elements in close proximity.

There is also natural variation in the humidity and temperature sensors used in the instrument. Measurement of any change in these environmental variables are crucial for producing accurate predictions from MOS sensors. As our calibration is unique to each sensor any offset produced through manufacturing variation should be compensated for, and not affect predictions of gas concentration. Although not germane to the central theme of measuring NO<sub>2</sub> and O<sub>3</sub>, the work done on characterising humidity and temperature variation is described in an appendix **B**.

# 5.1 Active airflow - yes or no?

Instruments can be designed to use fans to cycle air through a vented enclosure. Alternatively they can have an enclosure that carries the sensor elements on the outside, exposed to the atmosphere. Which approach is best is not a settled question among commercially available instruments[217].

Two prototype instruments (SOGS version 0.2.5) were placed in different casings. One of these enclosures was equipped with fans; the other had a cutaway that allowed the surface of the sensors to be directly exposed to the air beneath the casing. These were connected to the same power supply in parallel, and installed 25 centimeters apart from each other on the AURN station for calibration over a period of twelve days. The data used in this section was taken between the 19th and 28th of December 2014. The experimental setup is shown in figure 5.1. The sensors were both set to take measurements every five seconds. In the fanequipped casing, a measurement was immediately preceded by a second-long pulse of the fans.

The work described in this section was carried out in collaboration with A. P. Brundle (Brundle [235]), and first described in a paper written by the author (Peterson et al. [202]).



FIGURE 5.1: Experimental setup for investigating aspirated (on the left) vs. passive sampling (right side), on the AURN station.

#### 5.1.1 Results

Figure 5.2 shows the timeseries and scatter of predicted data from both sensors against the AURN reference from this brief experiment. Since only one side of the sensors was exposed to the atmosphere for the passive airflow instrument, only the NO<sub>2</sub> predictions from those exposed sensors were are used. Visually the passive airflow device appears markedly less sensitive, although it responds to a significant increase in background NO<sub>2</sub> late on the 27th of December. The statistics for the predictions in this experiment are shown in table 5.1 on the following page - note that these statistics are taken during the calibration period and are not directly comparable with the statistics over predictions outside it, such as those in section 3.8 on page 143. *FE* and *RSE* for the passive instrument are close to twice that of the active airflow sensor.



FIGURE 5.2: Timeseries and scatter of active airflow instrument (red) and passive airflow instrument (blue). The timeseries plots these predictions alongside the AURN reference (black). The graph on the right is a scatter plot of the same data in the same colours, with the 1:1 line marked as a dashed black line.

		RSE $\mu g m^{-3}$	$\mathbf{FE}$	PCC
NO2	Active	6.97	0.21	0.94
	Passive	13.37	0.35	0.76

TABLE 5.1: Table of performance statistics for the passive and active airflow instruments compared with the reference over the entire experimental period.

#### 5.1.2 Discussion

The timeseries in figure 5.2 appears to show that the passive sensor responds to changes in atmospheric  $NO_2$ , but is unable to resolve  $NO_2$  spikes that occur over a timescale of only a few hours. This can be modelled as the passive airflow causing a sensor to report accurate concentrations, but averaged over a period. To determine the equivalent integration time of the passive airflow sensor, a moving average was applied to the predictions from the active airflow instrument. The larger the window width of the moving average, the worse the quality statistics for the active sensor become. The equivalent integration period for the passive sensor is the window width at which the statistics for the two instrument configurations crossed. The results of this procedure are shown in figure 5.3. The RSE of the active instrument data becomes worse than the passive instrument at an integration time of 20 hours, for FE the crossover is around 24 hours, and for PCC the crossover is around 21 hours. The period varies between performance metrics, but overall is consistent with a window of roughly a day.



FIGURE 5.3: Graphs of change in performance metrics for the active and passive instruments, when the active airflow instrument has a moving average applied to it. The period of the moving average window is shown in the horizontal axis, and ranges from zero (ie. no averaging window) to thirty-six hours. The horizontal dashed black line is the passive instrument's performance statistics.

#### 5.1.3 Summary of findings

The instrument exhibits a marked difference in sensitivity when fans are not in place. It is clear from figure 5.2 that under the conditions of our experiment, passive airflow instruments sensors are not useful for hourly measurements. Passive airflow is acceptable for experiments that require only one measurement a day, but much of the advantage of the high time-sensitivity of MOS sensors in general is lost in this scenario. Using a fan in an instrument design is a simple precaution that can vastly improve sensor performance.

# 5.2 Warm-up experiments

The time taken for the MOS sensors to reach chemical equilibrium after a change in their power condition was investigated. When the sensor elements are cold, many atmospheric gases condense onto their surface. These gases inhibit sensitivity to the target gas by taking up binding sites on the sensor element. Bringing the semiconductor up to temperature will gradually evaporate them away.

The sensor's heating element draws a significant amount of power. There is an incentive to have it powered down, or at least put in a low power state, whenever practical for battery-powered mobile sensors. The strategy chosen must be compromise, because changing the power supply to the MOS heater has been shown to adversely effect performance[228].

Beyond a deliberate strategy as part of the experiment design, there are several means by which a sensor deployed in the field might have its power interrupted momentarily. For example, if there is no battery backup, installing the sensors in a car means disconnecting their power supply for up to five minutes before they are reconnected to the car battery. During one mobile campaign a sensor, equipped with a transceiver, experienced periodic brownouts when the transceiver tried to connect to the GSM network to deliver data every six minutes or so. This caused periodic spikes in the voltages from the MOS sensors.

Equation 3.5 is the pseudo-static equation giving the response of the sensor at a point in time. The a terms are proportional to the actual concentration of gas in the atmosphere, only when the rates of the equilibrium reactions are all faster than the rate at which the atmospheric composition in front of them varies. In situations where the sensor element cools to room temperature, the above can no longer be assumed, as the Arrhenius law reminds us that reaction rate decreases exponentially as temperature falls[236].

The net rate at which airborne molecules of a particular gas adhere to the sensor surface can be expressed as:

$$\frac{da_{gas}}{dt} = [G_{gas}]\lambda_a(T) - \lambda_d(T)a_{gas} - (A - \sum_{i \in G} a_i)$$
(5.1)

which is the sum of the positive attachment of the gas to the sensor and the negative separation of gas from the sensor surface.  $\lambda$  is a constant representing the likelihood of attaching and detaching respectively, as a function of surface temperature T. a is the area of the sensor that is taken up by attached gas molecules either of the gas in question or for i all other gases, A is the surface area of the sensor. From equation 3.6, the area already taken up by other gases is an issue if the interaction lifetime of the gases with the sensor become significant on the timescale of the experiment. This results in sensor hysteresis and non-zero response time during normal operation.

In a static atmosphere after a long enough period of time, in other words in steady state conditions, equation 5.1 will settle to zero. An equivalent equation for a for all gases is:

$$\frac{da}{dt} = \sum_{i \in G} \frac{da_i}{dt} = \left(\sum_{i \in G} [G_i] \lambda_{ai}(T) a_i - \lambda_{di}(T) a_i\right) - (A - a)$$
(5.2)

and it is the  $a_i$  term that appears in  $f_i(a_i([G_i]), T_a, T_h)$  in equation 3.5. Consequently, solving this partial differential equation can give a baseline level of gas adsorption that the sensor will tend to on a long enough timescale. Due to the non-linearity of f this will not help us in practical terms, and the baseline must instead be found experimentally.

Alternatively, if power is interrupted, the surface temperature will change and the  $\lambda$  terms will change with it. It is hypothesized that this will cause the amount of gas adsorption to move toward the new baseline following an exponential curve:

$$a(t) = k_a e^{(\lambda_a - \lambda_d)t} + k_c \tag{5.3}$$

where the constant  $k_c$  represents the change in voltage between powered and unpowered states once the sensor has reached equilibrium.  $k_a$  is a constant that must be found experimentally, and t is the time since the power was turned off. The time constant of this decay (the time in seconds for the voltage to change by 63% of the way between the powered or unpowered baseline) is equal to  $\lambda_a - \lambda_d$ ).

With this hypothesis, a power supply interruption episode will produce a sensor output voltage jump between before, and after the interruption with the characteristic form shown in figure 5.4. The longer the sensor is left unpowered, the larger the effect. This hypothetical model is based on the unphysical assumption that the temperature will instantly change when power is interrupted. In reality the  $k_c$  term will change as the temperature of the surface varies over time. This may also contribute to a hysteresis effect, where the actual voltage change will deviate from the exponential curve shortly after the power is connected or interrupted.

When the power supply is cut and restored, there will be a point at which the contribution to the sensor's resistance as a result of the power interruption event is minimal compared to other sources of error within the sensor. This is the amount of time until the sensor can be said to be functioning normally again. If the sensor obeys the laws described above, then the amount of time before the sensor is ready
for operation will be predictable, and it will vary with the duration of the power interruption.



FIGURE 5.4: Example curve of net gas adsorption after a power supply interruption event.

# 5.2.1 Experimental method

The instrument cannot take data when the power is disconnected, so a method was devised to allow the behaviour of the sensor voltage while unpowered to be characterised. A MOS instrument was placed in a closed plastic container in a dark corner of an air conditioned lab for three days to approach a stable atmosphere. A number of tests were run whereby the instrument's power supply was interrupted for a set duration, and when the sensor was repowered the shape of the voltage curve afterward was recorded as the sensor surface returned to its powered equilibrium. Because the sensor takes several measurements a second, and the exact moment of those measurements cannot be externally determined, a strategy had to be made to ensure that the voltage increase during disconnection could be accurately recorded.

This consisted of two extrapolations of the sensor data. The first was before the disconnection period. The first, a linear extrapolation from the last recorded data point of the sensor to the beginning of the time of disconnection. The second, an exponential curve was fitted to the voltages after the power was restored, and extrapolated back to the end of the disconnection period. The difference in projected

voltage at either end of the power interruption is a single data point corresponding to the duration of that interruption.

This process was repeated with interruptions lasting from a second to several minutes. A time series diagram of the order and duration of these interruptions is shown in figure 5.5. Plotting the voltage gap versus interruption time should produce a set of points corresponding to the exponential curve in equation 5.3.

# 5.2.2 Results and discussion



FIGURE 5.5: A timeline of the power supply interruption experiment, which took place over multiple days in December 2016. The labels for the periods during which the power supply was interrupted are referenced in the results section.

The voltage responses from the four sensor elements in the instrument are shown in 5.6. The voltage curve after sensor disconnection has the characteristic form predicted by the model illustrated in figure 5.4 after the power connection was restored. The magnitude of the voltage jump after the power supply interruption, once hysteresis is compensated for, follows the exponential decay law described in equation 5.3. The statistics for this decay are given in table 5.2 on the following page

Concor	Baseline	$k_a$	Decay time
Sensor	shift $k_c$ (V)		constant (s)
OXa	$0.465 \pm 0.031$	0.405	549
OXb	$0.619 \pm 0.035$	0.548	745
O3a	$0.242 \pm 0.026$	0.210	423
O3b	$0.209 \pm 0.025$	0.182	340

TABLE 5.2: Table of statistics for the exponential voltage decay curve following power interruption for the four sensors. Constants are defined in equation 5.3.

Immediately after the supply was reconnected the oxidising gas sensors showed increasing voltages for a couple of seconds before they began decaying down to the powered baseline level. These indicate sensor hysteresis that was more pronounced for the oxidising gas sensors than the ozone sensors. The presence of sensor hysteresis is suggested by the assumptions in the model. While it is important to account for this during analysis, it has no bearing on the key conclusions of this experiment.

# 5.2.3 Summary

As shown in table 5.2, different magnitudes of response  $(k_a \text{ and } k_c)$  were observed for different sensors, despite the ambient conditions being the same. This points to the inherent differences between sensors due to the manufacturing process. Consequently the exact form of the decay should be expected to be individual to each sensor.

Strictly speaking the time until the sensors are usable will depend on when the change in output voltage from the power interruption effect diminishes to below the instrument voltage noise. Unfortunately the results given here cannot be converted into absolute concentrations because the sensor used was not recently calibrated.

Time constants for the decay range from just under six minutes to just over twelve. The time constant is lower (and hence the delay before the sensor becomes usable



### Power supply interruption

FIGURE 5.6: Fitted curves after power supply interruption for four different sensors, two each of oxidising (OXa and OXb) and ozone (O3a and O3b), in a stable atmosphere. The vertical axis of each graph is the increase in voltage either side of the disconnection period. The baseline that the sensor output settled to when left powered for at least an hour was set at zero volts (see figure 5.5. The blue curves are the raw voltage curves, some of which show the hysteresis of the sensor toward the beginning each curve (particularly for OXa and OXb). The red lines are continuations of the voltage curves to the time at which the power supply was reconnected, ignoring the hysteresis. These terminate in red circles at the projected voltage difference at the moment the sensor was reconnected. The dashed green line is an exponential curve fitted to these points showing the likely voltage curve when the power is disconnected and the sensor is cooling. The solid green line is the unpowered baseline voltage inferred from this fit. The grey lines either side of the likely voltage curve delineate the 95% confidence interval of the exponential fit.

again is shorter) for the ozone sensors. This is not an impractical amount of time. Waiting an hour should be several times the length of the decay constant no matter the sensor. In field experiments with mobile sensors, an hour's delay after power is restored has no major effect on results.

# 5.3 Manufacturing variation in MOS sensors

Metal Oxide Semiconductor gas sensors exhibit a high degree of individuality due to variations in the manufacturing process[139]. The unique architecture of the MOS sensor board used in the instrument prototype, with two sensors of the same kind separated by a centimeter and shielded from sunlight, permits the true level of correlation between sensors to be calculated.

# 5.3.1 Method

This analysis spans every single experimental dataset taken with the two-sided MOS sensor board (see figure 3.4). Because the sensors are on either side of a PCB, they are exposed to almost the same atmosphere. There are 173 such datasets spanning four years of experiments, but to ensure a fair comparison, the sensors must be fully active and ready to take measurements. Consequently the list of datasets was trimmed to remove the first hour after power-up as per section 5.2.3 on page 188, and only datasets with at least ten records after trimming were included in the statistics. The list was also trimmed of duplicate sets of data.

To prevent saturation issues caused by the amplifier electronics reaching the edge of their voltage range (see figure 6.4 in section 6.1 on page 196) - an error due to limitations in the electronics rather than flaws with the sensors - voltages below the 5th percentile and above the 95th were also trimmed from the datasets before they were compared with each other. This left 136 datasets. Datasets that pass this criteria were aggregated, and the statistics relating sensor a to sensor b were collated - *PCC* between these two sensors, as well as the slope and intercept of their voltages. The resulting statistics, in aggregate, are represented in figure 5.7, and in table 5.3 on the following page. This data is representative of the sensors used in the last four years with the Earth Observation Science group at the University of Leicester's physics department, but each data point is from a different *experiment*, and Earth Observation Science colleagues may have used the same sensor multiple times.

The output voltage of the sensor electronics for this experiment is linearly proportional to the element resistance, unless the sensor saturates.



Comparison between a/b sensor outputs

FIGURE 5.7: a) Violin plot of *PCC* between a and b sensors. The mean correlation has been marked with a white dot. b) Gradient of least-squares fit between a and b sensor voltages, with a value of one implying a 1:1 correlation. Violin plots are broader where more results are in the region indicated by the vertical axis.

PCC		Percentile		Number of sensors:		
Sensor	Mean	25th	50th	75th	-0.1 to 0.1	Less than $0.1$
OX	0.88	0.89	0.98	0.99	1	4
RED	0.58	0.28	0.82	0.97	14	13
O3	0.80	0.92	0.98	0.99	1	10
NH3	0.67	0.66	0.91	0.98	17	4

TABLE 5.3: Table of *PCC* - mean and percentile - between sensor pairs.

### 5.3.2 Discussion

The PCC between sensors is an important statistic. The sensors all will have individual baseline resistances and sensitivities that caused them to respond to atmospheric gas by a different magnitude, but these differences can be compensated for during calibration. The PCC, on the other hand, is high only if the pair of sensor outputs move in the same direction in response to a given change in the atmospheric composition. Assuming the same composition of atmosphere is in front of each sensor, PCC should be equal to one in the ideal case. PCC of zero implies one sensor is nonfunctional, and a negative PCC implies one sensor out of the pair is producing completely aberrant results.

From 5.3, the oxidising gas (OX) and (O3) sensors have the most consistently high PCC. The narrow tails on the violin plot for OX sensors show that a tiny minority of devices do not exhibit strong positive correlation with each other, and only one out of the entire dataset had negative correlation (less than -0.1).

The O3 sensor violin has the best mean PCC out of any type of sensor, although it has a very thin tail that stretches into negative correlation, representing 10 out of 136 sensor pairs. Very few ozone sensors were entirely nonfunctional, with one from the entire dataset having PCC between -0.1 and 0.1.

The reducing gas sensor (RED), not used elsewhere in this thesis, gave a very interesting result. The PCC violin plot shows that there's a significant proportion of sensors that exhibited very poor PCC, and the gradient violin in figure 5.7 b) for

this type of sensor reveals that roughly half as many sensor pairs had a gradient of zero between them as a gradient of one. A zero gradient indicates that one sensor in the pair had output that was locked to the same value no matter what the other sensor's response was. This is a very high failure rate.

The ammonia (NH3) sensors exhibited a less severe version of the same problem that plagued the reducing gas sensors, where from 5.7 b) about a third of the sensor pairs had a zero gradient. The poor *PCC* caused by this issue is reflected in the statistics in table 5.3 on the preceding page. The median *PCC* for the RED sensors is an acceptable 0.82, but the lowest quartile results are 0.28 - downright terrible. Other sensor types showed relatively good *PCC* between sensor pairs even at the lowest quartile.

However, that is not to say that the reducing gas sensor is useless. The highest quartile of *PCC* performance is very good for all sensor types, including reducing gas sensors. If the top quarter of all sensor pairs are performant, this means that each MOS instrument board, with a pair of each sensor type, has a 50% change of at least one sensor being consistently functional.

For those sensor pairs that do not exhibit correlation between each other, further experiments in a closed gas cell might show more clearly what species the sensors are responding to, or whether a large number of them are completely unresponsive.

# 5.4 Summary

The manufacturing variations in MOS sensors are quite pronounced, but for the OX and O3 sensors used for the majority of this thesis experience good correlation with one another, with 75% of them having a *PCC* of 0.88 for OX sensors and 0.92 for O3 sensors. Once properly calibrated the variation is less of an issue -

except that unless the amplifier electronics for the sensors are carefully designed, saturation might well result.

This sheds some light on the problems experience in previous chapters with long term experiment sensors c and d (see sections 3.6 on page 133, 4.2.2 on page 152 and 4.3.2 on page 160). From the results in this chapter it is statistically likely that the OX and O3 sensors were working well, but were still unable to produce usable results. This implies the problem is environmental and might be related to poor ventilation for the c and d sensors in the enclosure rather than manufacturing error.

The warm-up time experiment found that the decay constant for the effects induced after an interruption in sensor power supply can be up to 12 minutes depending on the sensor (see table 5.2 on page 188), which indicates that a period of an hour is needed before any experiments are attempted after powering up.

The usefulness of fans in ensuring fresh air is cycled onto the sensors is also shown in this chapter, with fan-assisted MOS sensors tracking well with the one-hour AURN data as shown in figure 5.6.

# Chapter 6

# Experimental study on spatial variation of pollutants

The individual exposure of human beings to  $NO_2$  is determined by quite smallscale variation in the concentration of the gas. Around a building or on either side of a fence,  $NO_2$  can vary significantly and this level of granularity in pollutant concentration on the scale of meters has itself been the subject of multiple studies [237][238][141].

In addition, the reactions in the Leighton cycle takes place on a scale of seconds[44]. This rapid variation increases the difficulty in using highly sensitive instrument which can measure changes with similar rapidity, like those based on MOS sensors. To obtain typical concentrations of atmospheric pollutants, sensors need to be sampled multiple times and then averaged together. Even then, there is significant variability between measurement sites depending on the local geography[237]. The high time resolution of the instrument developed in this work presents new opportunities to investigate  $NO_2$  in detail.

The question we seek to answer in this chapter is this: Through using mobile MOS instruments combined with mapping data, what can we learn about the local atmosphere's spatial and temporal variation of NOx,  $NO_2$  and  $O_3$  concentrations? What can we learn about the performance and utility of our sensors for measuring personal exposure?

Two experiments are described in this chapter. The first is an "adjacency" experiment. A rack of adjacent sensors on the AURN were run for 14 days, and the variation between these sensors give clues as to the characteristic length and timescale of  $NO_2$  variation.

The second experiment is a mobile deployment, albeit a very local one. The University of Leicester AURN site is next to a small sports field bounded by two roads, and colleague Amrita Aujla pulled a trolley with two Zephyr-prototype instruments strapped to it around the field. The influence of the roads on this raw dataset can be precisely determined thanks to the high time resolution of the sensors.

# 6.1 The arm's length scale

The calibration rig on the AURN station has fittings and power connectors to host up to ten sensors at once. Between the 26th September and 12th October 2016, nine sensors were set running side-by-side, along with the long term experiment previously analysed in chapter 3. While three of these sensors were found to be unreliable, the remaining six took a fortnight's worth of data in this way.

The purpose of this section is determine the variation in space of  $NO_2$  and  $O_3$  concentrations, in the environment near the Leicester AURN station. Since the Zephyr instrument's architecture has two identical sensors less than a centimeter away from each other in each housing, it is possible to determine a baseline level of sensor inaccuracy by looking at the difference between these pairs of sensors which should be receiving air of the same chemical composition at the same time.

From there, the variation of  $NO_2$  in the atmosphere may be determined by looking at the variation between adjacent instruments, which are distributed over a couple of meters.



# 6.1.1 Experimental setup

FIGURE 6.1: Layout of the multi-sensor adjacency experiment. Blue lines indicate dimensions. Positions are given relative to the leftmost sensor's intake. The instruments (apart from the LTE, long term experiment) are demarcated by a name of the form DCAxx, where xx are two numerals. The red crosses indicate sensors that were found to not produce accurate results - see later in this section.

Figure 6.1 shows the layout of the sensors on the side of the AURN station. The sensors were set to take measurements every ten seconds, although the LTE sensor (not used in this experiment) instead takes measurements every 20 seconds. The AURN station reports data every hour (although this data is the average of much more rapid measurements). While the sensors took measurements every ten seconds, they are not necessarily perfectly time-synchronized; to allow comparison between them, the timestamps of the sensors in the following analysis are averaged to the nearest whole minute.

Critically, the sensors are shielded from the direct emissions of the nearby University Road by being on the far side of the AURN building. They are all roughly 20 meters from the road, but proximity to a road has been noted as a significant determining factor in average pollution exposure[141]. The airflow around the building was not modelled for this experiment.

This analysis uses a subset of data from the period between the 27th of September and 11th October 2016, starting at 12 noon and lasting two weeks.

As depicted in figure 6.1, some of the instruments did not perform well enough for this experiment and are omitted from further analysis. Specifically, while the two pairs of OX sensors in DCA88 and DCA79 agreed with each other well enough, they had terrible correlation with the other sensors in the experiment, particularly DCA88. These sensors also gave trouble with the humidity and temperature sensors - this implies a problem with either the sensor ventilation or electronics. These sensor outputs are omitted from future analysis, along with DCA41, which had a damaged temperature sensor.

Within each sensor are two pairs of four different instrument sets, identified from now on as a and b - such that the sensor a inside DCA66 is referred to as "DCA66a".

# 6.1.2 Voltage correlation analysis

Even raw voltages from these sensors provide interesting information. A set of scatters from all the used sensors is shown in figure 6.2.



Grid of scatter plots for adjacent O3\OX sensors

FIGURE 6.2: This scatter grid plot shows the correlation from all six functional sensors (that is, twenty four sensor elements, twelve each for OX and O3 MOS sensors). The bottom left portion of the graph shows the O3 sensors, the top right are OX. Pairs of a and b sensors are adjacent to each other and are shaded the same colour, helping emphasize the blocks of four scatter plots. The axes are normalised so that the center of each plot axis is the mean voltage of the sensor for that axis. The scale is individual to each scatter, and is such that both axes of the scatter cover a range of voltage equal to twice the sum of standard deviations for both of the sensors indicated by the axes.

The results from section 5.3 on page 190 are borne out here in the comparison between sensors in the same housing. The only case where the PCC between sensors in the same instrument fell below 0.99 was in the case of DCA82, where the O3 sensor reported a decent PCC of 0.97 and a suspiciously poor PCC of 0.92 for the OX sensor.

Aside from this, the shape of the voltage correlation between different instruments should be noted. The scatter plots are each normalised to the sum of standard deviations of both sensor axes, and where the slant of the scatter deviates from 45° it implies a lower sensitivity for one of the sensors. For both of the O3 sensors in DCA52 versus DCA46, the scatter shows a consistent slant within each instrument - a narrow range of the DCA46 sensor voltage corresponds to a large range of DCA52 voltage. This pattern is repeated elsewhere, with sensors within an instrument exhibiting similar slants to each other compared to other instruments. As the slant is per instrument, rather than per sensor, this implies the cause is the change in environment between each housing, rather than due to manufacturing variations between sensors.

Some sensors exhibit saturation - that is, their response is curtailed by the limited range of voltages that the amplifier electronics can produce. The DCA82 OXb sensor is a clear example. For the lower part of each scatter for the OXb sensor, there is a flatly vertical line. This represents a range of values where the sensor does not respond to changes in the environment that are detected by the sensor it is being compared to.

### 6.1.3 Fitting analysis

For both  $NO_2$  and  $O_3$ , fits were taken using the inverse equations against the reference instruments on the AURN station. The calibration was produced from

the entire two week dataset of the experiment. The prediction used the faster (minute-averaged) data from the Zephyr instruments and covered the same period.

In the same way as earlier, the predicted timeseries data was compared between sensors and instruments. The results are shown as a grid in figure 6.3. The takeaways are similar to those of figure 6.2, in that most of the sensors in the same enclosure correlate well to each other in their predictions. Table 6.1 on the following page shows a summary of the correlations of the sensors with each other, and with the reference instruments.

As might be expected, on average the sensors within the same instrument housing correlate best with each other, followed by their neighbouring instruments, and poorest with the AURN reference instruments. Because the prediction is occurring during the same period as the calibration, the AURN statistics cannot be directly compared with the results from chapter 3 (which test predictions outside the calibration period), but they were not appreciably worse. The confidence intervals for fractional error between other instruments are very large, particularly for  $O_3$  - the variation between sensors is such that the error bar is two and a half times larger than the average result.

Some sensors performed noticeably worse than others. The sensors in DCA45 gave predictions with a PCC of 0.75 between its sensors, but the other scatters involving DCA45a evidence poorer performance in general, with a consistently shape in the scatter graph (figure 6.3. DCA82 is still poor for both a and b sensors in NO<sub>2</sub>, with a clear cutoff where the OX sensors were saturated visible in the scatter plots.

	Comparison	$ m RSE~\mu gm^{-3}$	$\mathrm{FE}$	PCC
	a vs b	$12.0 \pm 9.7$	$0.38\pm0.35$	$0.89 \pm 0.16$
NO2	vs other sensors	$13.0 \pm 8.8$	$0.73\pm0.99$	$0.88\pm0.18$
	vs reference	$21.7\pm6.6$	$0.40\pm0.19$	$0.76 \pm 0.18$
	a vs b	$5.7\pm3.4$	$0.17 \pm 0.19$	$0.99\pm0.01$
O3	vs other sensors	$9.48\pm 6.3$	$0.80\pm2.02$	$0.96\pm0.05$
	vs reference	$8.3\pm2.6$	$0.47 \pm 0.27$	$0.97\pm0.02$

TABLE 6.1: Table of average correlation of the sensor predictions between each other, divided into two categories: sensor a vs sensor b from the same instrument, and the sensors compared with different instruments. Error ranges are 95% confidence interval. The graph also has statistics for how well on average the instruments correlate with the AURN reference, using hour-averaged data from the Zephyr prototype instruments.

# 6.1.4 Residuals and atmospheric variation

In this section a model with some significant assumptions will be constructed, and used along with data from the Zephyr instruments to calculate the length scale over which  $NO_2$  and  $O_3$  concentrations change in the atmosphere.

This "variation" of the gas in space is comprised of the change due to emission, deposition, and any chemical reactions that produce or consume the gas. The typical variation observed between two sensors will depend crucially on wind direction and how pollution sources in the region are carried to the sensor, as well as (in the case of the Leighton gases whose equilibrium is moderated by the presence of UV photons) the diurnal cycle on a longer timescale. Effects like plumes of fresh pollution cannot be represented by a simple Gaussian noise distribution with a characteristic length. However, for the purpose of this analysis, we will model the atmospheric gas concentrations in these terms. The distribution of a gas then, when measured at a single point, has a characteristic time related to the wind speed and the rate at which chemical reactions occur.

In this model, the average change between measurements will equal zero overall over timescales much larger than this characteristic time. The  $NO_2$  characteristic diurnal cycle occurs over a much shorter period than the timescale of this



Grid of scatter plots for O<sub>3</sub>\NO<sub>2</sub> predictions

FIGURE 6.3: This scatter grid plot, similar to 6.2, shows the correlation between the predictions of the twenty four sensors - in the upper right for  $NO_2$ , in the lower left for  $O_3$ . As with figure 6.2, the scale of the individual scatter plots centers around the mean of the predicted data and extends to the sum of the standard deviations. This keeps the data centered but does not transform it or change the angle of the individual plots. Colour is used to visually group a and b sensors from the same instrument together. These scatter plots are between

Zephyr-prototype sensors, not the AURN reference instruments.

experiment (two weeks), and so this analysis will proceed disregarding the diurnal effect.

The difference in the concentration of a gas [G] at a point in time, as measured by two separate sensors i and j, is the sum of the inherent detection error of each sensor  $\delta S_i$  and  $\delta S_j$  in its finite capacity to measure the actual value of the target gas, and the actual difference in gas between the two sensors  $\Delta[G_{ij}]$ .

If the overall change in gas during the course of an experiment is zero, then the standard deviation of the local change in gas between two of these sensors is equal to the *RSE* between them. The standard atmospheric variation in gas for a particular sensor  $\Delta[G_i]$ , over a length scale l, can then be expressed as the average of all  $\Delta[G_{ij}]$  terms except those where i = j (which are the same sensor) and  $i = j \pm 1$  (which are the adjacent sensors in the same instrument housing). This gives twelve results for each gas, one per sensor, and the average of these values is the standard atmospheric variation among all the sensors - that is, the magnitude of the gaussian noise model for the sensors over a characteristic length scale l.

The 95% uncertainty in the standard atmospheric variation estimate is equal to

$$2 \times \sum_{i}^{N} \sqrt{\delta S_{i}^{2} + \sigma[G_{i}]^{2}} / \sqrt{N}$$

$$(6.1)$$

where N is the number of sensors and  $\sigma[G_i]$  is the standard deviation of the differences in gas concentration between sensor i and all other sensors:

$$\sigma[G_i] = \sqrt{\sum_{j \neq i}^N \sqrt{\Delta[G_{ij}]}/N} \tag{6.2}$$

If the measurement error of any sensor is Gaussian, it can be approximated with the reference residual, which is a measure of the error between the calibrated sensor and the reference instrument. The AURN reference instrument itself has a fractional uncertainty of better than 14% for NOx and related gases[239, 87]. The *reference residual* is a measure of the extent to which the predicted data differs from the reference. It can be calculated as:

$$\Delta[G_s] = \sqrt{\frac{\sum_{i \in N} \left( [G_{Ref}]_i - [G_s]_i \right)^2}{N - 1}}$$
(6.3)

where N is the number of data points, [G] is the concentration of the target gas,  $\Delta[G_s]$  is the error of the gas concentration for that individual sensor,  $[G_s]$  is the sensor's reported gas concentration and  $[G_{Ref}]$  is the reference instrument's concentration. There are twelve sensor elements, and this gives us twelve results for each target gas. In order to compute the residual, the sensor data must be averaged to the hourly results given by the AURN.

The average of the reference residual for all of the sensors is  $11.19 \pm 3.12 \,\mu g \,m^{-3}$  for NO<sub>2</sub> and  $3.97 \pm 1.24 \,\mu g \,m^{-3}$  for O<sub>3</sub>. The mean value gives a typical uncertainty for sensors of this kind under these calibration conditions between their prediction and the reference, although for the rest of this analysis the individual error values calculated in equation 6.3 for each sensor are used as error bounds.



FIGURE 6.4: Residual between sensors for  $NO_2$ . Each point is the standard deviation of the residual standard error between the labelled sensor and all the others in the experiment. The x axis arranges the instruments in their approximate position on the AURN station, with results for sensors a and b for each instrument (in red and blue) shown next to the intra-instrument comparison between a and b(in grey). The separate axis on the left of the diagram shows the average of the inter-instrument comparison. The error bars are to the first standard deviation (63% confidence).



FIGURE 6.5: As above, but for  $O_3$ .

The residual of each sensor between different instruments  $\sigma[G_i]$ , and for those sensors are shown in figures 6.4 and 6.5. For NO<sub>2</sub>, the large error bars from individual instruments are clearly shown to the the biggest source of error for this calculation - an error margin that the averaging between sensors will go some way toward compensating for. For O<sub>3</sub> the factional error is lower, consistent with results obtained throughout this work. While the intra-sensor results are inside the margin for error for the adjacent instruments, they are also generally lower - average difference between the other instruments and the paired a and b sensors is

 $0.50 \pm 7.75 \,\mu g \, m^{-3}$  for NO<sub>2</sub>, and

 $1.89 \pm 3.86 \,\mu g \,m^{-3}$  for O<sub>3</sub>, with error bounds at the 95% confidence level. That the sensors within the same casing have a lower standard deviation is to be expected, but the margin for error is very high.

For NO<sub>2</sub>, the standard atmospheric variation averaged over all sensors is  $6.5 \pm 7.28 \,\mu g \, m^{-3}$ .

For  $O_3$ , it is  $4.73 \pm 2.75 \,\mu g \,m^{-3}$ , to 95% confidence.

The characteristic length scale of the modelled gaussian distribution l, for this setup, is the average of the sensor separations in the experiment - equal to  $0.92 \pm 0.68 \,\mathrm{m}$ .

## 6.1.5 Summary

Through this experiment, the length scale of the variation in NO<sub>2</sub> and O<sub>3</sub> was calculated. It is difficult to find a comparable figure to this experiment from the literature, because the analysis requires sensor-to-sensor RSE to be calculated, not just the residual against an reference. Surveys featuring a large number of instruments left near the same reference station for a long period would be suitable if the raw data from these could be acquired. The studies from Denver[50] (6-8 sensors), or Oslo[185] (24) would be suitable.

The signal to noise ratio of these results is almost 1:1 for  $NO_2$ , although  $O_3$  gives a clearer prediction with 2:1 signal/noise. The greater margin for error on the NO<sub>2</sub> results might be due to the greater number of variables going into the determination of NO<sub>2</sub> concentration. Whereas the O3 sensor by itself can be fitted to reference  $O_3$  concentrations, the crosssensitivity of the oxidizing gas sensor means that it takes two of the MOS sensors to predict NO<sub>2</sub> concentration. Critically, each MOS sensor comes with its own set of unpredictable cross-sensitivities, or as figure 6.2 shows, potential electronic issues in the form of saturation. Having two sensor inputs into the equation rather than one can make a significant difference, above and beyond any innate difficulty in measuring atmospheric NO<sub>2</sub>.

While successful in determining a length-scale for these forms of airborne pollution, this section leaves open an important question. Since there is a worse correlation between adjacent instruments than with sensors in the same instrument, this implies that even over a range of half a meter or so the dynamics of the atmosphere and chemical changes among pollutants can affect the concentration of sensible gases significantly. But the same issue affects the correlation between the Zephyr prototype instruments and the AURN instruments they are calibrating against. How much of the error when calibrating Zephyr prototypes in this environment is due to the different detection method within the reference and the Zephyr, and how much is down to the distance between the AURN station inlet and the sensor intake fan? The sensors are capable of measuring atmospheric pollutants to a respectable accuracy using the calibration methods described in chapter 3, but the current calibration setup, with the air gap between sensors under test and the reference instrument intake, has serious room for improvement. Future work involving these sensors should mean a redesign of the calibration setup.

# 6.2 The garden scale

The examination of sensor response in the immediate environment of an atmospheric reference base station is not the primary purpose of developing small, portable sensors. To learn more about a population's exposure to air pollution, we must deploy them in more varied settings.

The experiment described in this section is a fairly straightforward one - build a mobile sensor, and see how the concentrations of gas change across a field. The objective is to examine the effects of road proximity, and the level of sophistication needed to understand how that proximity can be modelled. How much of the variation in NOx detected by the sensors is local (that is, produced by the nearby roads), and how much is deposited in the vicinity of the experiment through atmospheric transport?

# 6.2.1 Method

Between December 2015 and into early 2016, several experiments were performed by colleague Amrita Aujla with a roving sensor platform. The data for this analysis was taken on the 11th of December 2015. A cart was assembled with a long vertical pole and fitted with two Zephyr prototype sensors, one held 20 cm and one at 150 cm from the ground, as shown in figure 6.6. This cart was pulled by hand around the sports field adjacent to the AURN station and nearby roads and carparks, tracking the carts position using GPS data. The sensors took one data point every five seconds, recording MOS sensor voltages, temperature, humidity and time. GPS location information was recorded on a mobile phone GPS tracker (at the time, the sensor's onboard GPS was not yet available). For all the cart experiments, the sensors were turned on in the lab the night before and the cart was left outside for an hour to achieve equilibrium at the AURN station before being moved.



FIGURE 6.6: Roving instrument cart beside the Leicester University AURN station. Only the instrument at 150 cm elevation is used in this analysis.

The instrument at 20 cm closer to the ground was used by Aujla to examine the short-lived chemical species emitted by the ground. Unfortunately after calibration at the AURN station this instrument returned wildly inaccurate predictions. Calibration is carried out two meters from the ground, and it's possible that interference from other gases threw off the sensor predictions. The results from this sensor are not relevant to the following discussion - the measurements in this section come from the sensor at head height (150 cm).

The instruments were mounted on the AURN (not on the cart) for calibration both before and after the larger series of experiments performed by Aujla, giving two calibration periods - one between the 19th and 27th of November 2015, the second from the 21st of January to the 5th of February in 2016. A transitional fit would have been ideal to interpolate between these two periods, but the earlier period produced poor fits and so the second period was used. The statistics comparing the prediction from the sensor over the calibration period is given in table 6.2. These statistics cannot be directly compared with the averages shown in table 3.4 on page 134, because latter are averages of predictions that were made during a validation period that did not coincide with the calibration period.

		RSE	FE	PCC
NO2	a	14.27	0.344	0.87
	b	11.81	0.298	0.92
O3	a	5.24	0.05	0.99
	b	6.85	0.15	0.98

TABLE 6.2: Calibration statistics for the roving sensor. RSE values are in micro grams per meter cubed.

The cart was pulled around the small sports field next to the AURN station between 4:17pm and 4:31pm, in the afternoon of the 11th of December. The field is elevated from the pavement, with an embankment about a meter high, near the junction between Welford and University Road, bordered by a scattering of trees that go some way to shielding it from traffic. The experiment lasted fifteen minutes, but there were other experiments performed with this setup and the larger dataset not discussed here will be published elsewhere. The measurement interval for the roving sensor was five seconds.

## 6.2.2 Results

The predicted concentrations for both NO<sub>2</sub> and O<sub>3</sub> from the second calibration period, matched to their GPS location, are shown in figure 6.7. During the same parts of the path, the paired sensors returned similar concentrations, as evidenced by similar colours on the diagram. This is explicit in figures 6.8(a) and 6.8(b) show scatter plots comparing the predictions of the a and b sets of sensors (*not* between the sensors and the AURN reference). Subfigure a) shows NO<sub>2</sub>, for which the predictions exhibited a *RSE* of 1.38 µg m<sup>-3</sup> and a *PCC* of 0.99. Subfigure b) shows O<sub>3</sub>, which had a *RSE* of 0.81 µg m<sup>-3</sup> and *PCC* of 0.99. While the sensor



FIGURE 6.7: Path and relative concentrations of the roving sensor experiment superimposed on an aerial image of the ground near the AURN station. The average distance moved between measurements is  $6.93 \pm 2.05$  m.

predictions had diverged from each other slightly since calibration, they showed excellent PCC and very low residual standard error.

The looping path described in figure 6.7 does not show a very strong correlation between road proximity and the amount of NO<sub>2</sub> measured. Instead, figure 6.9 shows a steady decline in NO<sub>2</sub> concentration as time progresses. Two things make it likely that this is not due to degradation or otherwise aberrant performance



FIGURE 6.8: Scatter comparing a and b sensor predictions during the mobile sensor experiment. Predictions are made using the end calibration period.

from the sensors: First, the two sensors in the instrument correlated well with each other, and second the dips in  $O_3$  concentration (particularly noticeable early on in the timeseries) anti-correlate with  $NO_2$ , which is consistent with the idea that these gases are linked together by the Leighton relationship and responding to elevated  $NO_2$  in the environment due to traffic pollution. A more detailed analysis on the influence of the pollution from the road is needed to clarify the extent to which correlation between the sensors proximity to it and the amount of  $NO_2$  and  $O_3$  it detects should be expected.



FIGURE 6.9: Timeseries of predicted concentrations, for both  $NO_2$  and  $O_3$ .

Gaussian modelling[240] (unrelated to the Gaussian noise model of section 6.1.5 on page 207) is a popular component in simulations of the movement of pollution in the atmosphere. A central assumption is that pollution emitted from a source forms a plume that is carried by the wind and spreads as it dissipates. The simplest Gaussian model neglects deposition, turbulence or atmospheric chemical reactions, and uses wind direction to simulate a plume from a source of pollution. The plume equation provides quite a mathematically simple and computationally fast means for determining concentration c downwind from a source[241, 3]:

$$c(x,y,z) = \frac{Q}{2\pi\sigma_y\sigma_z u} exp\left(\frac{-y^2}{2\sigma_y^2}\right) \left(exp\left(\frac{-(z-h)^2}{2\sigma_z}\right) + exp\left(\frac{-(z+h)^2}{2\sigma_z}\right)\right)$$
(6.4)

where x is downwind, y is crosswind, and z is vertically upward. Q is the source term related to the amount of the emission per unit volume, and the  $\sigma_y$  and  $\sigma_z$ terms indicate the crosswind and vertical mixing. u is the wind speed and h is the release height of the source of emission. The last term of this equation represents reflection from the ground.

The plume modelled in this equation does not change along the downwind (x) axis, based on the assumption that when the wind speed is high, transport is more dominant than diffusion. However when this is not the case then the modelling for dispersion of the gas becomes much more complex, at least when an absolute prediction of plume density is desired. In practical terms knowing the "extent" of a plume from the source is difficult as it depends on the definition of when the plume has completely dissipated - generally when the airborne concentration in a plume reaches some fraction of the concentration outside the pollutant source [242].

If a time and wind independent pollution distribution is assumed - a fairly bold assumption - then quantifying the effects of a nearby road can be greatly simplified. In the case of a single point source at distance r from an instrument:

$$c(r) \propto r^{-p} \tag{6.5}$$

where for a three-dimensional diffusion p is equal to 3. In case of a linear source, the proximity factor becomes:

$$c(r) \propto \int_{L} r(l)^{-p} dl \tag{6.6}$$

where r is the distance from each point along the road L. An array of road proximity data can be used to express this as a dimensionless, normalised quantity by dividing the collected factors c by the minimum value of the array.

To model the linear source of pollution from the road, the simulated path of Welford Road and University Road was divided into 1 meter increments, and the distances from each section point were summed together, performing a numerical integration of equation 6.6. A scatter plot of this information is shown in figure 6.10. For this calculation, the different length of the longitude and latitudinal degrees at  $52^{\circ}$ N was accounted for.

There is a very poor correlation between road proximity as measured above and  $NO_2$  concentration observed for both sensors; if the hypothesis is true then we should see a clear straight line on the graph for  $NO_2$  going diagonally to the top right. The anti-correlation with  $O_3$  is to be expected, and suggests that the sensors were operating properly. The assumptions made for this analysis may be too audacious to produce reliable results in these conditions - particularly neglecting the effects of wind transport.

The alternative is to replace the simple arithmetic with computational fluid dynamical (CFD) modelling that takes wind speed and direction into account, matching the predicted exposure from the road with NO<sub>2</sub> concentration.

The wind measurements on the AURN take place every minute, and the roving sensor measurements every five seconds. To match the two, the roving sensor data points were matched to the wind reading with the nearest timestamp, re-using neighbouring wind measurements if needed.



FIGURE 6.10: Scatter of predicted concentration vs. road proximity for  $NO_2$  sensors a and b. The road proximity factor is a normalized, dimensionless figure obtained by dividing the road proximity equation (6.6) by its minimum value.

A set of CFD data was generated based on the nearby buildings, with their geometry measured using LIDAR (light detection and ranging) by Infoterra, courtesy of Leicester City Council. The dataset was the same one used by A. JeanJean for modelling urban air quality[84]. Each data file was specific to a wind speed and direction and the set covered 360° of wind direction in 10° increments, for wind speeds of 2, 4, 6 and 8 meters per second. Each data file consisted of a number of unsorted vertex definitions on a 10 meter square grid of vertices, excluding points that would have been inside structures.

Each data point from the experiment was matched to the nearest data files using two rounds of linear interpolation - one between the wind speeds and directions represented by the different CFD files, and one rounded to match the GPS position data from the roving sensor to the nearest vertices within those files.

During the experiment, the wind speed generally stayed below 2 meters per second, and the wind mostly blew from the west, from the direction of University Road. For experimentally measured wind speeds lower than the lowest speed described in the CFD dataset, the lowest values (those at 2 meters per second) were used rather than interpolating between that and zero gas exposure. Where the wind reading was zero, those data points were omitted completely from this analysis. A wind scatter plot for the experiment is shown in figures 6.11(a) and 6.11(b), focusing on the limited set of wind speeds experienced during the experiment. The anticorrelation between NO<sub>2</sub> and O<sub>3</sub> is visible, but there is no clear link between wind direction and concentration of either gas.



FIGURE 6.11: Scatter of wind direction during roving sensor experiment, coloured according to a) NO<sub>2</sub> or b) O<sub>3</sub> concentration from sensor a within the 150 cm instrument. The colour indicates the gas concentration measured by the sensor for a particular wind direction. A random number (between -0.05 and +0.05 m/s) was added to the x and y direction of each wind measurement to make distinct the different concentration in groups of data points.

# 6.2.3 Discussion

The results of this analysis are shown in figure 6.12. They are scarcely more encouraging for the idea that there should be a strong correlation between road proximity and pollutant concentration than those for the simpler model. The left side of the graph, representing road influence of less than 4, has points covering the full span of predicted concentrations. There are horizontal clusters of points around  $47.5 \,\mu g \,m^{-3}$  and  $36 \,\mu g \,m^{-3}$ , which means that closer to the road these concentrations were detected consistently. While the upper patch approaches higher road influence, it is a very poor correlation overall.

The *PCC* between modelled influence and measured gas is 0.122 and 0.096 for the NO<sub>2</sub> fits, and -0.176 and -0.166 for O<sub>3</sub>. Again, the anti-correlation between NO<sub>2</sub> and O<sub>3</sub> that demonstrates a link between independent sensor readings and the

atmospheric chemistry is present. Overall the data shows that the changes in reported gas concentration in these experimental conditions appear to be dominated by other factors than road proximity.



FIGURE 6.12: Scatter graph of simulated road influence compared to measured  $NO_2$  and  $O_3$  concentrations of sensor a. Note the marginal influence of the road, for the given levels of influence from the road.

One possible explanation is that transportation of pollutant gases from elsewhere in the city might dwarf the pollution produced by traffic on University Road.

As discussed in section 1.4.1 on page 31, the NO<sub>2</sub> in the urban environments can be either emitted or generated photochemically in the immediate vicinity, advected from nearby local sources, or imported from distant regions by PAN transport. Thanks to the Leighton cycle, locally produced concentrations of NO<sub>2</sub> and O<sub>3</sub> are dependent on the NOx concentration - this is not the case with NO<sub>2</sub> that has been deposited from elsewhere in the atmosphere. Figure 6.13(a) is an example from [89], showing a graph of oxidant gases (the sum of NO<sub>2</sub> and O<sub>3</sub>) and NOx, which is the sum of  $NO_2$  and NO. The regional contribution labelled varies seasonally [89], although some events like fires can elevate regional production and transport. Using plots like these to examine the AURN data can reveal how much influence PAN transport has on the local  $NO_2$  concentration.

Figure 6.13(b) is the plot of oxidant verses NOx gas for the University of Leicester AURN station in the month of December 2015. A high proportion of the detected pollution is from regional sources, and the amount of this detected in the sports field will be independent of the geometry of the nearby roads. Note however that the hourly variation of Ox in this the AURN data, which has a standard deviation of  $11.17 \,\mu g \,m^{-3}$ . This is the uncertainty of the regional influence cutoff point, and is of comparable magnitude to the gradient of the trendline, which means that from day to day the proportions might vary, and road-based pollution could sometimes be significant.



Oxidant vs NOx graph for AURN station in December 2015



FIGURE 6.13: a) is a scatter showing of oxidant gases verses NOx taken at a six urban and suburban sites in and around London, averaged over the nighttime hours, during two Novembers in 1998 and 1999. Picture credit to [89]. The values are in part per billion, the translation of which to micrograms per meter cubed depends on the molecular weight of the gases in question. b) shows the scatter graph of oxidant gases (NO<sub>2</sub> plus O<sub>3</sub>) verses NOx taken at the University of Leicester AURN station, hourly, during November in 2015. The trendline and y intercept are marked with a solid and dashed line respectively.

The fitted trendline has the equation [Ox] = 0.353[NOx] + 65.54.
We can in this light re-examine the timeseries for the roving sensor experiment, plotting the NO<sub>2</sub>, O<sub>3</sub> and total oxidant concentrations back-to-back, alongside the CFD-simulated contribution of NO<sub>2</sub> emissions from the road, in figure 6.14. The grey line representing the simulated road influence bears very little relation to the movement of the sensor predictions for either gas. The poor *PCC* between the road influence and either target gases (0.09 and 0.12 for the NO<sub>2</sub> sensors, and -0.18 and -0.17 for the O<sub>3</sub> sensors) backs this up. This is in line with the measured oxidant levels, which average 73.18 and 72.76 µg m<sup>-3</sup> for sensor a and b respectively during the cart experiment, only slightly above the regional influence cutoff point  $65.59 \pm 11.17 \,\mu g \, m^{-3}$ , implying the bulk of variation in oxidant gases comes from non-local sources, and not the nearby roads.



FIGURE 6.14: Timeseries showing separate readings for sensors a and b throughout the roving sensor experiment, for each of the measured gases, and the total oxidant level, compared with the CFD simulated contribution of gas coming from the road, a dimensionless concentration factor.

### 6.3 Summary

In this chapter, two different scales of variation for NO<sub>2</sub> and O<sub>3</sub> pollution were examined. In the first experiment, it was found that the spatial variation of NO<sub>2</sub> was  $6.5 \pm 7.28 \,\mu g \, m^{-3}$  over a characteristic length scale of  $0.92 \pm 0.68 \, m$ . For O<sub>3</sub>, it is  $4.73 \pm 2.75 \,\mu g \, m^{-3}$  over the same scale.

In the second experiment,  $NO_2$  and  $O_3$  concentrations measured by a mobile sensor did not correlate well with road proximity. After using Gaussian modelling and then a CFD model to better measure the road proximity, the analysis in 6.1.5 on page 208 showed that, for the immediate surroundings of where the experiment took place, a modest result should not be unexpected. But these two case studies show the value of a low-cost gas sensor system that can take multiple measurements in quick succession in learning new things about the atmosphere.

A strength of small sensor systems is in their mobility. The roving sensor experiment described here gave a result that, while not immediately intuitive, is in line with what would be expected given a full consideration of the environment. Small sensors are able to determine pollution hotspots on a scale far smaller than is possible through measuring with the expensive static instrument suite of the AURN network, and although the results presented here are limited, they are a necessary step in developing robust procedures for measuring pollutant gases with instruments of this kind.

# Chapter 7

# Conclusions and further work

The increasing public awareness of the dangers of urban pollution is leading to changes in government policy in the UK, and a demand for more accurate measurement systems. The advantages of small sensors in cost and mobility, in spite of their mixed record for reliability compared to expensive chemiluminescence instruments, is leading to their increasing adoption as a useful supplement to traditional measurement systems.

It is in this context that this work was carried out. The literature regarding the usefulness of various small sensors in limited campaigns is quite rich at the time of writing, but there is a need not only for determining the accuracy of such devices, but also to establish proper techniques for calibration and deployment that can help improve reliability and deliver more useful results.

In this thesis, a bespoke sensor system is described in detail through development to initial calibration to novel field experiments. The best practices for atmospheric instrument design are examined in the context of measuring  $NO_2$  pollution, but they are more broadly applicable to any instrument that samples the atmosphere outdoors. These principles are realized in the SOGS platform, the flexibility of which is already lending itself to commercial applications such as the Earthsense Zephyr<sup>[243]</sup>.

Although not the focus of this doctorate, the author played a major role in designing, building and testing the physical hardware of the instruments used. This included circuit layouts and control software, as well as the data analysis software briefly described in chapter 3. The author produced a paper[202] describing preliminary results from the instruments, alongside many of the findings related in chapter 5 of this thesis.

The calibration of metal oxide semiconductor gas sensors is examined at length. The evolution of the fitting equations is tracked from the simple relations in section 3.2.2 on page 99 used to compare performance of the sensors in a closed cell with a spectrometer reference instrument, to a detailed comparison of several alternative equations in section 3.4.1 on page 117. The equations that worked best for the two target gases were:

$$[NO_2] = k_c + \frac{k_1}{V_{OX}} + \frac{k_2}{V_{OX}}RH + \frac{k_3}{V_{OX}}T + \frac{k_2}{V_{OX}V_{O3}}RH + \frac{k_2}{V_{OX}V_{O3}}T$$
(7.1)

for  $NO_2$  and

$$[O_3] = k_c + \frac{k_1}{V_{O3}} + \frac{k_2}{V_{O3}}RH + \frac{k_3}{V_{O3}}T$$
(7.2)

for  $O_3$ , where the  $k_i$  terms are constants that are found during fitting of the sensor output voltages  $V_{O3}$  and  $V_{OX}$  to the reference concentrations. *RH* and *T* are the output of the instrument's humidity and temperature sensor respectively. It was determined that calibration in the environment next to a reference instrument, rather than in a carefully controlled gas cell, were likely to produce more realistic results.

Optimal calibration time for outdoor calibration was shown to be roughly a week in section 3.4.2 on page 120, with longer calibrations not necessarily producing better predictions. The novel CAFPOLD data analysis technique developed by the candidate during this thesis work (see section 3.4.3 on page 124) provides a way to quantify changes in how a sensor performs over time through long-term positioning at calibration sites. Elsewhere in the literature, rolling calibrations are mentioned, but CAF-POLD allows for the performance of a sensor's technology to be evaluated in a very general sense. Using this analysis, the average performance of a sensor over the course of one month from calibration is shown to be:

	$RSE \ \mu g \ m^{-3}$	FE	PCC
$NO_2$	$23.6 \pm 3.9$	$0.67\pm0.27$	$0.63 \pm 0.05$
$O_3$	$23.3 \pm 3.9$	$1.11\pm0.65$	$0.80 \pm 0.04$

Simple fitting is not always enough to ensure reliable results from instruments containing these sensors, and it is shown in section 3.7 on page 137 that selecting good calibrations by checking their predictions immediately after the calibration period can help with this, at the cost of occasionally requiring longer calibration times. The tests are simple thresholds above which the statistics of the prediction must be for the calibration to be considered worth using. Several thresholds were tried out, and the most practically useful were a residual standard error of  $25 \,\mu g \, m^{-3}$  for NO<sub>2</sub> and a fractional error of less than 0.5 for O<sub>3</sub>. This technique achieved predictions over the month from calibration with the following statistics:

	$RSE \ \mu g \ m^{-3}$	FE	PCC
$NO_2$	$22.9 \pm 5.3$	$0.80 \pm 0.44$	$0.57 \pm 0.13$
$O_3$	$20.0 \pm 5.05$	$0.57 \pm 0.43$	$0.85\pm0.05$

The usefulness of transitional fits, particularly for roving sensors, was demonstrated in section 4.3.2 on page 160. These fits require two separate calibrations on either side of an experimental period, but on average were shown to increase performance by:

	$RSE \ \mu g \ m^{-3}$	FE	PCC
$NO_2$	$+3.44 \pm 3.00$	$+0.10 \pm 0.05$	$-0.08 \pm 0.09$
$O_3$	$+1.72 \pm 1.85$	$+0.05 \pm 0.03$	$+0.03 \pm 0.03$

The record of MOS gas sensors has been mixed when tested by other groups, occasionally delivering very poor results for unclear reasons. Thanks to the unique paired-sensor equipment used in this work, the correlation between sensors and their manufacturing variability can be demonstrated on a large scale, as in section 5.3 on page 190. For example, 19% of reducing gas MOS sensors exhibited zero gradient compared with their partner in the same instrument. In spite of this, the two sensors in the same housing have better correlation and lower error between their predictions than sensors in neighbouring housings, as demonstrated in section 6.1 on page 196:

	Comparison	$RSE \ \mu g \ m^{-3}$	FE	PCC
	a vs b	12.0 + 9.7	0.38 + 0.35	0.89 + 0.16
$NO_2$	vs other sensors	13.0 + 8.8	0.73 + 0.99	0.88 + 0.18
	vs reference	21.7 + 6.6	0.40 + 0.19	0.76 + 0.18
	a vs b	5.7 + 3.4	0.17 + 0.19	0.99 + 0.01
O <sub>3</sub>	vs other sensors	9.48 + 6.3	0.80 + 2.02	0.96 + 0.05
	vs reference	8.3 + 2.6	0.47 + 0.27	0.97 + 0.02

The question of airflow in sensor designs is addressed in section 5.1 on page 179, with the conclusion that for practical air quality monitors, a fan is an essential design feature. Sensors that didn't have a fan but were instead directly exposed to the atmosphere produced results consonant with an integration time of one day, which is a major performance bottleneck with MOS sensors, which are otherwise quite capable of returning a result every second.

In section 5.2 on page 183, the time after power-up before MOS sensors become usable is shown to be at least one hour.

A study is made of sensor performance in the context of changing wind direction and the different chemistry at different times of day. While the wind carrying plumes of interfering pollutants is shown to be a clear issue that affects sensors in a consistent way, the effects of day and night chemistry on the fidelity of MOS sensors are not pronounced, and might not justify the longer calibration time necessary to compensate for the diurnal cycle.

The lessons learned in this work point toward several avenues of future research. First, the concentration of NO<sub>2</sub> gas in the atmosphere can change quite rapidly through both atmospheric transport of local emissions and the Leighton cycle, and this is one of the reasons NOx concentration, which is preserved in the Leighton reactions, is a common marker for atmospheric pollution instead of NO<sub>2</sub>. No specific MOS sensor has been designed for measuring NOx. However, a combination of sensors may well be capable of fixing NOx concentration, and more besides - preliminary, albeit unpublished, results using an isokinetic gas sampling input split between an FTIR and a SOGS-Zephyr instrument calibrated the four types of MOS sensors equipped on our instrument for  $CH_4$  and  $N_2O$  response, using a linear gas equation (7.3) over four days.

$$[Gas] = k_c + k_1 T + k_2 R H +$$

$$k_o V_{OX} + k_o t V_{OX} T + k_o h V_{OX} H +$$

$$k_r V_{RED} + k_{rt} V_{RED} T + k_{rh} V_{RED} H +$$

$$k_z V_{O3} + k_{zt} V_{O3} T + k_{zh} V_{O3} H +$$

$$k_o V_{NH3} + k_o t V_{NH3} T + k_o h V_{NH3} H$$

$$(7.3)$$

where the  $V_{OX}$ ,  $V_{OX}$ ,  $V_{OX}$  and  $V_{OX}$  are the amplified voltages from the oxidising, reducing, ozone and ammonia sensors respectively, T and RH are temperature in kelvin and relative humidity respectively, and the k terms are fitting constants to be found. The results are shown in figure 7.1. Note that this prediction is over the same period as the calibration, but correlation is respectable even for this very preliminary experiment.



FIGURE 7.1: Timeseries of reference (from FTIR instrument) and MOS predicted gases,  $CH_4$  and  $N_2O$  respectively. The units are preserved from the FTIR - part per million and part per billion.

Oxidation of  $CH_4$  by OH radicals in the presence of NOx leads to tropospheric ozone production, in addition to being important to understand in its own right as a greenhouse gas and a byproduct of some hydraulic fracturing (fracking) techniques.

The high degree of variability in  $NO_2$  concentration in space and time underscores the need for good calibration setups. The effects of offsetting the timestamps of instruments during calibration have been documented in this thesis, and the atmospheric variability is described in field studies. To compensate for these effects, it might be worth rigging a calibration setup so that the same packet of air is split between reference instrument and calibrating sensor.

This atmospheric variability is also of key relevance to health and urban design. Further experiments using small, low-cost 3D anemometers on mobile sensors could contribute significantly to understanding the spatial and temporal variability of these important pollutants, from testing models to identifying pollutant hotspots to source attribution on a second-by-second basis. Sensors with UV-light detectors would allow the effects of inclement radiation on the Leighton cycle to be investigated.

Longer term experiments may be able to test models of PAN transportation and deposition. This transport has been noted as a significant source of urban pollution[238], and small sensors would be able to measure this transportation with a high degree of accuracy. Small instruments could be calibrated for  $NO_2$ ,  $O_3$  and NO, and possibly using electrochemical sensors if MOS sensors proved insufficient. The balance between these gases over a month-long experiment can indicate the contribution of local and regional gas sources to local pollution, but being able to measure this balance in multiple locations can give clues as to largescale pollution transport.

Frustratingly the MICS-2610  $O_3$  MOS sensor, which was vital to both the NO<sub>2</sub> and  $O_3$  predictions and performed better than the oxidising gas sensor, has been discontinued by the manufacturers following their takeover by another company. Similar sensor technology is not available elsewhere as of writing. It is fortunate that many of the techniques developed in this work are applicable elsewhere.

The limitations of a calibration procedure for such a rapidly varying and reactive gas species as  $NO_2$  that relies on an air gap between reference and sensor under test should be clear. In future work, care should be taken to ensure that the packet of air entering the setup is divided evenly, travels through the same conditions and reaches reference and sensor at the same time. This would likely permit shorter calibration periods and better predictions. Sadly such a setup was not achievable during this work.

Several other experiments done by the Leicester air quality group with mobile Zephyr prototype sensors have been made but are yet unpublished in the scientific literature. These take advantage of the key benefits of small sensor technology in hunting for pollution hotspots and investigating issues flagged up during computational simulation of the airflow within cities. Vehicular-mounted sensors, and sensors on bikes, are two scenarios that have been the subject of diligent work by colleagues of the author and a fuller treatment of the data is wise. In particular, best practices for ensuring representative sampling of air from vehicles have not been developed - isokinetic intakes that ensure even gas pressure through the instruments, no matter the speed of the vehicle, are almost certainly required to produce accurate data.

# Appendix A

# Technical Notes on the SOGS Instrument

## A.1 Introduction

This appendix describes the progression of the SOGS electronics over time. The features of this equipment evolved considerably during the course of this work, but the central objective has always been a modular bus for taking measurements from atmospheric pollution sensors in the field. The author engineered the circuit layout for SOGS and the instrument boards, and wrote all the firmware and computer software except the control functions for the GPRS transceiver.

The SOGS base board has at its heart a Master Control Unit (MCU), a small microprocessor. Supporting this is a power management circuit, and it in turn controls a set of Analog-to-digital converters (ADCs). These ADCs were connected to the amplified outputs of individual sensors, which on the oldest designs were fixed to the same circuit board as the MCU. Data logging involved a real time clock for timestamping and a data storage medium, and on some models a radio transceiver. The final version of SOGS (0.2.11c) that participated in the experiments described in chapter 6 was a very capable instrument platform. The SOGS base board contained a battery and solar power management circuit and a custom hardware interface for supporting several instruments at once. This allowed for both electrochemical and semiconductor sensors to be installed side by side on separate "instrument boards", backed up by humidity and temperature sensors on the base board. SOGS was designed to be cheap and easy to manufacture and deploy. The descendants of this SOGS instrument form the core of the commercially deployed "Zephyr", produced by Earthsense Ltd.

As the instrument was developed, more and more designed capabilities were enabled and tested. At the beginning of the research in this thesis the board design was at version 0.2.3, which mounted an XBee transceiver. During testing this proved very unreliable, and the SOGS board also had numerous circuit errors that made the battery charger unusable. During the project a new board (0.2.5) was produced with a modified design which omitted the radio and rebuilt the power circuits in a smaller form factor.

Throughout the development timescale, SOGS was used and tested as the core of several experiments carried out by other researchers, adding to the data value of performance.

### A.2 Revision history

#### A.2.1 Version 0.1

Only a single example of the first version of SOGS was made. This device was not fully modular, incorporating a bank of MOS sensors onto the main board that were fed through a 10-bit ADC. Power was supplied by inefficient linear regulators and controlled by an Arduino Nano daughterboard. This device was capable of 10-bit precision in its ADCs, and did not supply data for any of the work in this thesis, but nonetheless it was an important prototype as it demonstrated the sensitivity of the MOS sensors even at relatively low resolutions.

#### A.2.2 Version 0.2.3



FIGURE A.1: Version 0.2.3 layout. The Raspberry-Pi header is in the lower right. The major subsystems are identified by coloured blocks.

Version 0.1 was a limited test bed, but 0.2.3 used a different architecture and was built for experimental deployments. It was designed to be easy to hand-solder, and all ten SOGS were produced in this way in April 2014. The five main components of the board are:

- POWER: The board uses a switching regulator (to the right) to provide efficient DC power to the other components. The separate power segment on the left contained a shutdown circuit and linear regulator for backup purposes, but in this version of the board neither of them worked.
- MCU: The controlling computer for this board is an Arduino Nano 3.0, chosen for its wide support among the open source community. Custom software was written to interface with some components, as the default libraries were found to be non-functional.

- COMS: On this generation of the board, communication was provided through an XBee radio transmitter, which had an effective urban range of 100m. This was connected to the MCU through a four-way serial link that could also connect it to the Raspberry Pi header. The XBee did not work as advertised either on or off the board, and was dropped in future designs. This board version only communicated via a direct USB connection to the MCU.
- STORE: Data storage on an SD card requires accurate time stamping, so the board mounts a real time clock and battery. On this version of the board, the SD card was not tested due to a board error, but the clock was functional.
- SENSORS: This subsystem contains a 4-way multiplexer and a pair of 16-bit ADCs coupled to a programmable gain amplifier. The amplifier effectively allows measurements to be taken with improved resolution when they are within a nominal voltage range. The ADCs are capable of sampling a single bank of eight voltages at 16-bit resolution 3.5 times a second, and 40 times a second on 12-bit resolution. For sampling a single voltage, the sample rate can be as high as 160Hz. On this version, the sensors were limited to 15-bits of resolution.

The MOS sensor array attached to this version of SOGS differed from the manufacturer's reference design in two ways: First, the resistive element was sampled by a low bias current amplifier rather than being tied to the input voltage rail, which reduces passive heating, and secondly the amplifier circuit produced an output linearly proportional to the element resistance, both innovations intended to make modelling easier. The sensor amplifiers had adjustable gain settings via potentiometers, as MOS sensors vary greatly in base resistance from the factory and the required gains could not be predicted.

#### A.2.3 Version 0.2.5



FIGURE A.2: Version 0.2.5 layout. The major systems are identified by coloured blocks. The power subsystem is split into three locations on this design.

The first board designed for mass production, completed in August of 2014. Loss of the requirement for hand soldering meant that many components could be miniaturized - SOGS 0.2.5 is two thirds the size of 0.2.3, and the MOS sensors are one third the size of the previous generation. The wiring was optimized and many tracks were made significantly thicker to lower resistance, particularly to power supply circuitry. Naturally many errors were also corrected, opening the way to new systems being tested. Five boards were ordered from Newbury Electronics.

- Power: The new power supply architecture, aside from being much more functional, wasted less heat due to a better component selection and was much more compact. It contained a voltage-controlled switch that prevented issues with the slow warm-up time from the main power supply, intended to fix an issue raised during testing in Berlin. However, it still showed signs of unreliability in the long term.
- MCU: The MCU was unchanged.

- COMS: Gone was the troublesome XBee transmitter, and in its place a small header for attaching a Bluetooth slave board, which has been shown to be functional. The communication system between different components was also simplified.
- STORE: The SD card's fault was fixed and it had a new power supply to play with. These modifications lead to the SD card's full implementation as a means of storing experimental data.
- SENSORS: The sensors were upgraded to handle full 16-bit precision from their designed voltage range of 0-5V. Attempts to implement automatic scaling of the programmable gain amplifiers lead to spurious results in field tests, but the cause of this was isolated to the power system which was modified in the next version.

#### A.2.4 Version 0.2.11



FIGURE A.3: Version 0.2.11 layout

This board design is by far the most prolifically manufactured to date, with 20 models being produced by SMS Electronics who used a more controlled process. The newest sub-version (0.2.11c) had small improvements, replacing through-hole components with surface mount to reduce manufacturing costs; 50 instruments of this kind were produced.

- POWER: The power subsystem had better capacitor smoothing and a smaller loop for the switching regulator, reducing EMF noise. The better manufacturing process helped increase the reliability of this critical system. Many subsystems on the board were given individual power gate switches, including the sensor bank power supply, which could now be shut off for low power operation.
- MCU: Perhaps the most dramatic change on the board is to the MCU, which no longer lives on an Arduino daughterboard but is mounted directly to the PCB, increasing manufacturing control and reducing size and cost.
- COMMS: The COMMS subsystem worked well on the previous boards and was largely unchanged.

- STORE: The storage system also was not changed significantly.
- SENSORS: The sensor support electronics was compactified and put well clear of any power lines except the main supply line to reduce noise.

The board also hosted numerous corrections, some made possible due to the superior manufacturing technology employed at SMS, but space was also optimized and new features added. The form factor was kept the same for the sake of enclosure design, but new connectors for USB communication were added and the battery charge regulator circuit in particular was improved.

### A.3 Overview of most recent design

This version was used in all the experiments in this thesis aside from the Long Term Experiment (which used version 0.2.5). This section goes into detail on the components used and their design features.



#### A.3.1 Major component description

FIGURE A.4: Schematic diagram of SOGS 0.2.11c. The signals and power wires are represented by lines coloured to indicate their function. The grey boxes are subsystems that may be comprised of multiple electronic components. Circles are terminals that can accept single external connections or testing probes, rounded bars are headers that accept multiple connections. Pairs of half circles are links that can be soldered over to change the circuit's functionality. A dashed outline indicates an external component.

- MCU: The MCU is an 8-bit ATMega328. It has direct control over almost all of the multiplexers, and can switch on power to various subsystems to improve battery life. The MCU has three major connections: a 3x2 pin header used exclusively for reflashing new firmware, a mini USB-B socket which can be used for serial communication with a PC. It could also be used for firmware updates, provided a bootloader has been burned onto the MCU via the 3x2 header. A 5x1 pin header carried USB signals to a USB socket on the casing.
- SD card: The board features a micro-SD card slot (labelled uSD in figure A.4. A custom designed filesystem allows very high data density and fidelity, but means that the card could not be removed and put in a computer to retrieve experiments - instead a SOGS board must spool data via USB or GSM connection before it can be analysed. The filesystem uses "header" blocks to record the start and end blocks of separate experiment files. Each file data block was validated using Cyclic Redundancy Checking, and the headers were stored in triplicate in different locations on the SD card, allowing data to be automatically recovered in case of a power failure. 8GB capacity SD cards were used in the instruments in this thesis, but sizes up to 32GB are supported.
- R-Pi header: This header is a 2x13 pin, 2.54 mm pitch socket, pin-compatible with the Raspberry Pi computer. The header provides a breakout for all the main data busses at 3.3V logic level, and any circuits connected through it must use 3.3V logic to avoid damaging sensitive components. The power to the header can be switched on and off by the MCU. This gave SOGS instruments a powerful modular design, in addition to being useful for diagnosing problems. Two in-house peripheral daughterboards that used this header were an XBee short-range radio and a SIM928 GPRS transceiver. The header would allow the system to be integrated into larger devices for

example, with appropriate software a Raspberry-Pi can control all functionality on the board, and take advantage of SOGS communication, storage and battery/solar power capabilities.

- Real time clock: A M41T83 alarm clock is present on the board along with a socket for a 3V 20mm lithium button cell to power it. This clock is used for two things: timestamping data when the board is operating alone, and waking the board up if the MCU's sleep function is enabled.
- Humidity/temperature sensor (Kelvin/%RH): The sensor is a CC2D25, operating in digital output mode.
- Battery charge regulator/Maximum power point tracker (BCR/MPPT): This is a fairly complex circuit that takes between 6V and 20V DC input and converts it to 5V and 3.3V DC that the board can use. In designs that did not use solar panels, the main external power input uses the solar panel terminals. This system is also able to charge a battery and can either use float voltages or constant current charging. A thermistor (labelled T sensor on the schematic) can be used to provide a safety cutoff when cell temperature exceeds 40 degrees. The battery charge regulator is compatible with Lithium polymer, lead-acid and NiMH chemistries, although only Pb and NiMH batteries were used on deployment. The Maximum Power Point Tracker (MPPT) automatically adjusts the battery charging current to a level that would maximise the efficiency of the solar panels.
- ADC: The board uses two MCP3428 ADCs to convert analog voltages from the instrument board sockets into digital signals that are retrieved through the I2C bus. The MCP3428 takes samples at either 12, 14 or 16 bits of precision. Acceptable voltages are 0-5V, and the ADC makes use of a 2.048V precision reference. The ADCs have a programmable gain pre-amplifier which is used by the SOGS software to achieve additional precision at medium

voltage levels. The ADCs have a noise floor of 50 microvolts, just below the smallest detectable voltage increment of 70 microvolts.

- Instrument board sockets: There are four sockets along the top of the board for connecting instruments. The sockets are 2x6 pin, 2.54mm pitch. There are two ground pins, two 5V power pins and the rest are used for analog signals. In the firmware, the sockets are referred to as banks and numbered from 0 to 3.
- Bluetooth header: This is a pin header specifically designed for the BT-2S BlueTooth-serial slave board. An additional connection may be made between the status signal on the slave board and this header, which will go high and raise the MCU's interrupt 1 when the board is ready. Power to the bluetooth is controlled by the MCU, or can be set permanently on with a link. The Bluetooth functionality was used by colleagues in other experiments but did not directly contribute to the work in this thesis.

#### A.3.2 Supporting component description

- Level shifter: 3.3V logic and 5V logic coexist on the SOGS board. Putting 5V into a 3.3V device may damage it, so logic signals are fed through level shifters to protect them. The 5V digital devices on the board are the MCU, humidity/temperature sensor, clock and ADCs. The 3.3V devices are the bluetooth adapter (which strangely enough takes 5V power), and a Raspberry-Pi, if one is connected. The level shifter design does not work for signals with a frequency of greater than 10MHz.
- Multiplexers: The board carries six MCP14052B multiplexers (labelled MUX), which can direct analog and digital voltages like a set of single pole, quadruple throw switches. Four of them handle the signals from the instrument

board sockets, and two route serial signals among the board's four serial components (RSW) although the latter can be disabled by closing the nearby link. The multiplexers can only be controlled by the MCU.

• Power supplies: There is a high-current DC-DC converter on the board (an LMZ12003) capable of providing up to three amps at 5V, supported by a set of three low dropout regulators (labelled LDO) for low power circuitry. At several points there are diagnostic voltmeters that measure the external power supply voltages for housekeeping purposes.

# Appendix B

# Variation in temperature and humidity sensors

When using MOS sensors, temperature and humidity information is vital. Temperature affects the sensor element's response to different gases[197], and humidity strongly affects the device's output by binding to its surface[193]. Beyond this such information is useful for weather monitoring in general.

The manufacturer's rated precision of the humidity and temperature sensors are 0.3 °C and 2 % RH respectively, but this says nothing about their accuracy. While the calibration equations will compensate for incorrect values from the temperature and humidity sensors so long as they are not spurious, some investigation should be made into the accuracy of their absolute values.

### **B.1** Experimental setup

Between the 26th September and 12th October 2016, ten Zephyr sensors were set running within two and a half meters of each other outside on the Leicester University AURN station (as detailed in section 6.1). These sensors had various ages (up to two years) and deployment histories prior to that point, but the temperature and humidity sensors aboard them had all operated within their required tolerances prior to the experiment.

These sensors produced data between every ten and twenty seconds. To allow for a general comparison, these were averaged together into minute-long periods. The resulting data series are then compared in three ways: first, the Pearson Correlation Coefficient between each sensor is compared. Second the average offset between the sensors is calculated by simply taking an average of the each data series, and returning the difference between these averaged values. Finally, the slope between sensors is calculated - a 1:1 slope on average means that a change in temperature induces the same response in the two sensors.

### B.2 Results

Three of the sensors of the nine had defective electronics (as discussed in section 6.1) and returned zero or maximal values for the temperature and humidity sensors for this experiment. These have been omitted from this analysis.

Figure B.1 shows the correlation between temperature and humidity sensors for the six instruments. While the temperature sensors all agree to within two degrees of each other, with humidity it's a different story, with DCA82 and DCA45 reporting large (up to 10%) offsets from the other sensors. The offsets of both sensors are quite far from the error values given in the manufacturer's datasheet, emphasizing the need for practical checks on given physical specifications. DCA82's sensor also produced the worst correlation for both humidity and temperature readings, and evidence for this can be seen in the DCA82 scatters that have a larger grey cloud of anomalous readings around the relatively linear coloured patch in the center of each graph.



FIGURE B.1: Grid of scatter plots between six sets of Zephyr temperature and humidity sensors. The lower left grids show scatters between humidity sensors in different instrument housings, the upper right grids are for temperature sensors. The *PCC* between pairs of sensors is written at the bottom of the grid axes. Above each grid square is the mean and 95% confidence of the difference between the X and Y sensor. The scale of the grids is normalised such that the average of the data set is in the center, and the upper and lower axis boundaries are at points equal to this mean plus twice the standard deviation of the data. The scale is identical among all temperature and humidity grids respectively.

Two sources of error contribute to the level of correlation between adjacent sensor readings for humidity and temperature: different sensitivity and error for individual sensors, and the natural variation in the environment over the couple of meters by which the sensors were separated. Of these, the environmental effect would become more pronounced for sensors that were further from each other. Figure B.2 shows a comparison between sensor separation and PCC for both humidity and temperature. The trendlines are drawn neglecting DCA82, which was consistently worse than the other sensors in terms of PCC and accuracy, but even including this sensor a clear slope is visible, with the maximum correlation achievable with different sensors decreasing with increasing distance. Even so, PCC only drops to 0.94 for the sensors two meters away from each other.



Scatter of PCC vs. sensor separation

FIGURE B.2: Plot of *PCC* between sensors verses distance between sensors being compared, for both humidity and temperature. The diamond markers are for comparisons involving DCA82, and are omitted from the calculation producing the trendline.

The analysis and fitting used in this thesis depended on the relative humidity and temperature. However, the fitting does not depend on the sensor outputs precisely matching the real environmental conditions, only that they vary in the same way when those conditions change. Linear offsets to the sensor response are compensated for during calibration for gas concentration. Even the weaker sensors in this analysis produced high correlation with each other, better than 0.98 for sensors that are directly adjacent. So while using the humidity/temperature data with these sensors might require extra calibration work, they should not be a major source of error for fitting pollutant gases, where NO<sub>2</sub> PCC could be as low as 0.62, and O<sub>3</sub> PCC as low as 0.79 for a fit over one month (see table 3.4).

# Appendix C

# Case study: Random forest machine learning

The random forest is a machine learning technique that has proved helpful to other groups working on air quality monitoring sensors [229]. The output of a random forest is the averaged conclusion of a large number of individual decision trees, where each tree guesses the output gas concentration based on whether various input variables are greater or less than critical values that are randomly selected for each tree.

A random forest is grown from a calibration dataset, consisting of a timeseries of reference data and a number of input variable timeseries, in the following way:

- First, for a randomly selected input variable, the reference data is subdivided into two sets, one greater and one less than a critical point in the input data.
- Then, those two new data sets are further subdivided in the same manner. The branches of the tree bifurcate onward for an arbitrary number of steps.

- Each resulting "leaf" dataset is assigned an output value consisting of the average of the reference data points corresponding to remaining input values within that leaf.
- This process is repeated, growing a "forest" of random trees.

To produce a predicted output data point, the point is passed through each tree in turn, with branches being selected according to the point's value for that particular variable, and the forest's worth of leaf values are averaged together to give an output.

Random forests are useful for categorizing noisy input data, as they can qualify the prediction of the most common output of the averaged decision trees with the the proportion of dissenting outputs, which gives a measure of confidence in that prediction. Random forests have the weakness that they can't extrapolate past the furthest extents of the input dataset. When an input datapoint is beyond the bounds of the training dataset, the forest will produce the same output for that point no matter how far beyond the boundary the point lies, all other things being equal. They're also more computationally expensive than predicting based on a limited input point. Still, they can approximate an unknown non-linear function very effectively.

### C.1 Demonstration

Figure C.1 shows CAFPOLD analyses for a homemade random forest function trained to produce fits for NO<sub>2</sub> and O<sub>3</sub>, with the same OX and O3 sensors as are used in our more conventional regression fits. The random forest did not produce a functional result for NO<sub>2</sub> measurements with two out of the four sensors. A comparison of the performance of this analysis with the linear regression model used earlier in the thesis is shown in tables C.1 and C.2.



FIGURE C.1: Results of CAFPOLD analysis using random forest to predict  $NO_2$  and  $O_3$  concentration.

			RSE		FE		PCC	
		Method	Mean	SD	Mean	SD	Mean	SD
	a	Regress	23.15	3.23	0.77	0.40	0.65	0.03
NO2		Forest	21.87	2.50	0.60	0.22	0.57	0.09
NO2	с	Regress	23.37	3.63	0.68	0.25	0.62	0.05
		Forest	21.28	2.09	0.56	0.20	0.58	0.09
	a	Regress	23.84	4.02	1.25	0.80	0.80	0.04
		Forest	23.58	3.48	0.91	0.29	0.79	0.04
	b	Regress	23.27	3.88	0.94	0.50	0.79	0.04
$\cap$ 2		Forest	24.04	3.47	0.92	0.31	0.77	0.03
05	с	Regress	22.37	3.72	0.99	0.55	0.81	0.04
		Forest	23.42	3.35	0.90	0.34	0.79	0.04
	d	Regress	23.52	4.04	1.22	0.74	0.80	0.05
		Forest	23.59	3.36	0.90	0.34	0.79	0.04

TABLE C.1: Comparison between random forest ("Forest") and linear regression ("Regress") performance, in the form of averaged statistics over one month from calibration. The linear regression used the "inverse" equations detailed in section 3.4.1.

			RSE		FE		PCC	
		Method	Mean	SD	Mean	SD	Mean	SD
	a	Regress	27.22	5.61	1.67	1.39	0.54	0.11
NO2		Forest	23.17	5.31	0.81	0.47	0.52	0.09
NO2	с	Regress	25.44	6.78	1.14	0.79	0.56	0.11
		Forest	22.18	5.31	0.81	0.47	0.52	0.09
	a	Regress	25.18	5.59	1.45	1.03	0.81	0.06
		Forest	26.48	6.11	1.43	1.33	0.74	0.10
O3	b	Regress	24.62	4.92	1.06	0.74	0.79	0.07
		Forest	25.55	5.84	1.21	1.07	0.77	0.06
	с	Regress	24.89	6.04	1.19	1.02	0.82	0.06
		Forest	25.33	5.95	1.22	1.10	0.77	0.08
	1	Regress	25.35	5.90	1.32	1.22	0.81	0.06
	a	Forest	25.26	6.04	1.17	1.03	0.77	0.08

TABLE C.2: As above, but over five months.

The random forest technique produced better results compared with the linear regression models used earlier in the thesis in most cases. Over the one month average, for both target gasses the RSE and FE were both significantly lower and with a narrower standard deviation, although PCC was slightly worse. For the five month average the improvements for using a random forest were dramatic for

 $\mathrm{NO}_2$  (up to 20% better RSE, for example, and comparable PCC), but mixed for  $\mathrm{O}_3.$ 

The strength of the random forest approach is in fitting one value to another when the relationship is unknown. The interference in the oxidising gas sensors when used to detect  $NO_2$  from  $O_3$  is an issue that the linear regression models solved by determining an "empirical equation", which was capable of producing respectable fits. The random forest, by contrast, produces predictions that match (to a degree) the relations between the target and interfering gases that grew the forest during the calibration period; the "true" equation underlying those relations does not have to be known for the forest to produce these results. The markedly better results for  $NO_2$  even five months out are likely a reflection of this advantage. The effect is not as dramatic for  $O_3$ , the sensor for which was only compensated with temperature and humidity, without the explicit gas interference experienced by the  $NO_2$  sensor.

# Bibliography

- L. G. Jaccia. New static models of the thermosphere and exosphere with empirical temperature profiles. Technical Report 313, Smithsonian Astrophysical Observatory, May 1970.
- G. Proelss. Physics of the Earth's Space Environment: An introduction.
   Springer science and business media, 2012. ISBN 978-3-642-97123-5.
- [3] Southwest research institute glossary: Geocorona, April 2018. URL http: //pluto.space.swri.edu/IMAGE/glossary/geocorona.html.
- [4] T. Sharp and E. Howell. Earth's atmosphere: Composition, climate and weather. Technical report, Future US. Inc., October 2017. URL https: //www.space.com/17683-earth-atmosphere.html.
- [5] R. B. Stull. An introduction to boundary layer meteorology. Kluwer academic publishers, 2009. ISBN 978-94-009-3027-8.
- [6] UK Parliament. Public health act (section 91), 1875. URL http://www.legislation.gov.uk/ukpga/Vict/38-39/55/enacted.
- [7] UK Parliament. Clean air act, 1956. URL http://www.legislation.gov. uk/ukpga/Eliz2/4-5/52/enacted.
- [8] M. Kampa and E. Castanas. Human health effects of air pollution. Environmental Pollution, 151:362–367, January 2008.

- [9] M. Lippmann. Health effects of ozone: A critical review. JAPCA, 39(5): 672–695, 1989.
- [10] Automatic urban and rural network, 2018. URL https://uk-air.defra. gov.uk/networks/network-info?view=aurn. Website portal.
- [11] J. H. Gilliam and E. S. Hall. Reference and equivalent methods used to measure national ambient air quality standards (naaqs) criteria air pollutants -Volume I. Technical Report EPA/600/R-16/139, Environmental Protection Agency, 2016.
- [12] Edited by R E Hester and R M Harrison. Air Quality in Urban Environments. RSC Publishing, 2009. ISBN 978-1-84755-907-4.
- [13] Amir Givati and Daniel Rosenfeld. Air pollution in the UK 2016. Technical report, Department for the Environment, Food and Rural Affairs, September 2017. URL https://uk-air.defra.gov.uk/library/annualreport/ index.
- [14] Pope III, C. Arden, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, and G. D. Thurston. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*, 287:1132–1141, 2002.
- [15] Amir Givati and Daniel Rosenfeld. Quantifying precipitation suppression due to air pollution. Journal of the American Meterological Society, 43: 1038–1056, July 2004.
- [16] D. Vione, V. Maurino, C. Minero, E. Pelizzetti, M. A. J. Harrison, R. Olariuc, and C. Arsenec. Photochemical reactions in the tropospheric aqueous phase and on particulate matter. *Chemical Society Reviews*, 35:441–453, 2006.
- [17] P. Taylor. DEFRA national statistics release: Air quality statistics in the UK 1987 to 2018. Technical report, Department for the Environment, Food and Rural Affairs, 2018. URL https://assets.publishing.service. gov.uk/government/uploads/system/uploads/attachment\_data/file/ 796887/Air\_Quality\_Statistics\_in\_the\_UK\_1987\_to\_2018.pdf.
- [18] Department of Science, Information Technology, and Innovation. Airborne particles - continuous monitoring. Technical report, Government of Queensland, Australia, 2017. URL https://www.qld.gov.au/\_\_data/assets/ pdf\_file/0024/68532/particle-posters-web.pdf.
- [19] Air Quality Expert Group. Methods for monitoring particulate concentrations. In *Particulate Matter in the UK*, chapter 5. Department for the Environment, Food and Rural Affairs, 2005. URL https://uk-air.defra.gov. uk/library/aqeg/publications.
- [20] D. Harrison. Assessment of UK AURN particulate matter monitoring equipment against the January 2010 guide to demonstration of equivalence. Technical report, Department for the Environment, Food and Rural Affairs, 2010. URL https://uk-air.defra.gov.uk/assets/documents/reports/ cat14/1101140842\_Assessment\_of\_UK\_AURN\_PM\_Equipment\_against\_ 2010\_GDE.pdf.
- [21] Thermo Scientific. Thermo Scientific 8500 FDMS Filter Dynamics Measurement System Specification Sheet, 2010. URL https: //assets.thermofisher.com/TFS-Assets/LSG/Specification-Sheets/ D19418~.pdf.
- [22] PPD42 sensor datasheet. Shinyei Technology, March 2016. URL http: //www.shinyei.co.jp/stc/optical/main\_ppd42ns\_e.html.

- [23] E. Swietlicki, S. Puri, and H. Hansson. Urban air pollution source apportionment using a combination of aerosol and gas monitoring techniques. *Atmospheric Environment*, 30:2795–2809, August 1996.
- [24] Air Quality Guidelines. World Health Organization, 2005. URL https: //www.who.int/phe/health\_topics/outdoorair/outdoorair\_aqg/en/.
- [25] J. L. Jimenez, M. R. Canagaratna, N. M. Donahue, A. S. H. Prevot, Q. Zhang, J. H. Kroll, P. F. DeCarlo, J. D. Allan, H. Coe, N. L. Ng, A. C. Aiken, K. S. Docherty, I. M. Ulbrich, A. P. Grieshop, A. L. Robinson, J. Duplissy, J. D. Smith, K. R. Wilson, V. A. Lanz, C. Hueglin, Y. L. Sun, J. Tian, A. Laaksonen, T. Raatikainen, J. Rautiainen, P. Vaattovaara, M. Ehn, M. Kulmala, J. M. Tomlinson, D. R. Collins, M. J. Cubison, E., J. Dunlea, J. A. Huffman, T. B. Onasch, M. R. Alfarra, P. I. Williams, K. Bower, Y. Kondo, J. Schneider, F. Drewnick, S. Borrmann, S. Weimer, K. Demerjian, D. Salcedo, L. Cottrell, R. Griffin, A. Takami, T. Miyoshi, S. Hatakeyama, A. Shimono, J. Y Sun, Y. M. Zhang, K. Dzepina, J. R. Kimmel, D. Sueper, J. T. Jayne, S. C. Herndon, A. M. Trimborn, L. R. Williams, E. C. Wood, A. M. Middlebrook, C. E. Kolb, U. Baltensperger, and D. R. Worsnop. Evolution of organic aerosols in the atmosphere. *Science*, 329: 1525–1529, December 2009.
- [26] EU Standard Methods for monitoring and UK Approach, 2018. URL https://uk-air.defra.gov.uk/networks/monitoring-methods?view= eu-standards.
- [27] D. Riordan and Farah Adeeb. Air quality monitoring for sulfur dioxide in metropolitan Adelaide. Technical report, Environment Protection Authority, October 2004.
- [28] National air quality objectives and European Directive limit and target values for the protection of human health. Department for the Environment, Food

and Rural Affairs, 2012. URL https://uk-air.defra.gov.uk/assets/ documents/Air\_Quality\_Objectives\_Update.pdf.

- [29] J. Peuelas and J. Llusi. Plant VOC emissions: making use of the unavoidable. Trends in Ecology and Evolution, 19:402–404, August 2004.
- [30] R. T. Lidster J. D. Lee M. J. Evans A. R. Rickard A. C. Lewis J. F. Hamilton R. E. Dunmore, J. R. Hopkins. Diesel-related hydrocarbons can dominate gas phase reactive carbon in megacities. *Atmospheric Chemistry and Physics*, 15(17):9983–9996, September 2015.
- [31] R. Atkinson. Atmospheric chemistry of VOCs and NOx. Atmospheric Environment, 34:2063–2101, 2000.
- [32] B. Turpin and J. Huntzicker. Identification of secondary organic aerosol episodes and quantitation of primary and secondary organic aerosol concentrations during SCAQS. *Atmospheric Environment*, 29:3527–3544, December 1995.
- [33] M. A. K. Khalil and R. A. Rasmussen. Sources, sinks, and seasonal cycles of atmospheric methane. *Journal of Geophysical Research*, 88(C9):5131–5144, June 1983.
- [34] G. Myhre, D. Shindell, F. M. Bron, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J. F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang. 2013: Anthropogenic and natural radiative forcing. (Climate change 2013: The physical science basis. contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change). Technical report, IPCC, 2013. URL http://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5\_Chapter08\_FINAL.pdf.

- [35] I. C. Prentice. The carbon cycle and atmospheric carbon dioxide. In TAR Climate Change 2001: The Scientific Basis, chapter 3. IPCC, 2001. URL https://www.ipcc.ch/site/assets/uploads/2018/02/TAR-03.pdf.
- [36] D. Archer, M. Eby, V. Brovkin, A. Ridgwell, L. Cao, U. Mikolajewicz, K. Caldeira, K. Matsumoto, G. Munhoven, A. Montenegro, and K. Tokos. Atmospheric lifetime of fossil fuel carbon dioxide. *Annu. Rev. Earth Planet Sci.*, 37:117–134, 2009.
- [37] K. Permentier, S. Vercammen, S. Soetaert, and C. Schellemans. Carbon dioxide poisoning: a literature review of an often forgotten cause of intoxication in the emergency department. *International Journal of Emergency Medicine*, 10(1):14, April 2017.
- [38] NASA Earth System Research Laboratory and Global Monitoring Division. Trends in atmospheric carbon dioxide, 2018. URL https://www.esrl. noaa.gov/gmd/ccgg/trends/global.html#global.
- [39] M. Diallo, B. Legras, E. Ray, A. Engel, and J. A. Ael. Global distribution of CO2 in the upper troposphere and stratosphere. *Atmos. Chem. Phys.*, 17: 3861–3878, 2017.
- [40] P. Y. Foucher, A. Chedin, R. Armante, C. Boone, C. Crevoisier, and P. Bernath. Carbon dioxide atmospheric vertical proles retrieved from space observation using ACE-FTS solar occultation instrument. *Atmos. Chem. Phys.*, 11:2455–2470, 2011.
- [41] J. D. Allan L. Bruhwiler P. Forster D. Fowler A. Lauer W. T. Morgan P. Paasonen M. Righi K. Sindelarova E. Von Schneidemesser, P. S. Monks. Chemistry and the linkages between air quality and climate change. *Chemical Reviews*, 115(10):3856–3879, 2015.

- [42] H. Levy II, P. S. Kasibhatla, W. J. Moxim, A. A. Klonecki, A. I. Hirsch, S. J. Oltmans, and W. L. Chameides. The global impact of human activity on tropospheric ozone. *Geophysical Research Letters*, 24:791–794, 1997.
- [43] S. Beirle, U. Platt, M. Wenig, and T. Wagner. Weekly cycle of NO2 by GOME measurements: a signature of anthropogenic sources. Atmospheric Chemistry and Physics, 3:2225–2232, 2003.
- [44] P A Leighton. *Photochemistry of Air Pollution*. Academic Press, New York, 1961. ISBN 9780323156455.
- [45] A. S. V. Shah, J. P. Langrish, H. Nair, D. A. McAllister, A. L. Hunter, K. Donaldson, D. E. Newby, and N. L. Mills. Global association of air pollution and heart failure: a systematic review and meta-analysis. *The Lancet*, 382, July 2013. URL http://www.sciencedirect.com/science/ article/pii/S0140673613608983.
- [46] N. Gruber and J. N. Galloway. An earth-system perspective of the global nitrogen cycle. *Nature*, 451:293–296, 2008.
- [47] J. W. Erisman, P. Grennfelt, and M. Sutton. Nitrogen emission and deposition: The european perspective. *Scientific World Journal*, 1:879–896, 2001.
- [48] J. N. Cape. Surface ozone concentrations and ecosystem health: Past trends and a guide to future projections. *Sci. Total Environment*, 400:257–269, 2008.
- [49] L. N. Lamsal, R. V. Martin, A. van Donkelaar, M. Steinbacher, E. A. Celarier, E. Bucsela, E. J. Dunlea, and J. P. Pinto. Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument. *Journal of Geophysical Research*, 113(D16), August 2008.

- [50] R. Piedrahita, Y. Xiang, N. Masson 1, J. Ortega, A. Collier, Y. Jiang, K. Li, R. P. Dick, Q. Lv, M. Hannigan, and L. Shang. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques*, 7:3325–3336, 2014.
- [51] N. Bruce, R. Perez-Padilla, and R. Albalak. Indoor air pollution in developing countries: a major environmental and public health challenge. Bulletin of the World Health Organization, 78(9):1088, January 2000. URL http://www.scielosp.org/scielo.php?script=sci\_arttext&pid= S0042-9686200000900004.
- [52] D. G. Fullerton, N. Bruce, and S. B. Gordon. Indoor air pollution from biomass fuel smoke is a major health concern in the developing world. *Tropical Medicine and Hygiene*, 102(9):843–851, September 2008. URL https: //academic.oup.com/trstmh/article-abstract/102/9/843/1887437.
- [53] G. McGill, L. Oyedele, and K. McAllister. An investigation of indoor air quality, thermal comfort and sick building syndrome symptoms in UK energy efficient homes. *Smart and Sustainable Built Environment*, 4(3):329–348, 2015.
- [54] A. P. Jones. Indoor air quality and health. Atmospheric Environment, 33: 4535–4564, 1999.
- [55] World Health Organization regional office for Europe. Indoor air quality research, volume 23. WHO, 1986. ISBN 92-890-1269-2. URL http://inis. iaea.org/search/search.aspx?orig\_q=RN:23022155.
- [56] B. Sahlberg, M. Gunnbjrnsdottir, A. Soon, R. Jogi, T. Gislason, G. Wieslander, C. Janson, and D. Norback. Airborne molds and bacteria, microbial volatile organic compounds (MVOC), plasticizers and formaldehyde in dwellings in three north european cities in relation to sick building syndrome (sbs). Science of the Total Environment, 444:433–440, February 2013.

- [57] D. Norbck, I. Michel, and J. Widstrm. Indoor air quality and personal factors related to the sick building syndrome. *Scandinavian Journal of Work*, *Environment and Health*, 16(2):121–128, 1990.
- [58] J. A. Skinner, K. A. Lewis, K. S. Bardon, P. Tucker, J. A. Catt, and B. J. Chambers. An overview of the environmental impact of agriculture in the UK. Journal of Environmental Management, 50:111–128, 1997.
- [59] J. P. Majra and N. Mazzeo. Air quality in rural areas. In Chemistry, Emission Control, Radioactive Pollution and Indoor Air Quality, chapter 23. InTech, October 2011. URL http://www.intechopen.com/books/ chemistry-emission-control-radioactive-pollution-and-indoor-air-quality/ air-quality-in-rural-areas.
- [60] X. Pan, Y. Kanaya, H. Tanimoto, S. Inomata, Z. Wang, S. Kudo, and I. Uno. Examining the major contributors of ozone pollution in a rural area of the Yangtze river delta region during harvest season. *Atmos. Chem. Phys*, 15: 6101–6111, 2015.
- [61] M. Ninneman, K. L. Demerjian, and J. J. Schwab. Ozone production efficiencies at rural New York state locations: Relationship to oxides of nitrogen concentrations. J. G. R.: Atmospheres, 124:2363–2376, 2019.
- [62] S. Henninger. Biogenic isoprene and its impact on human health in dependence on meteorological conditions. *Journal of Environmental Protection*, 3: 1206–1212, August 2012.
- [63] J. Kesselmeier and M. Staudt. Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. *Journal of Atmospheric Chemistry*, 33:23–88, 1999.
- [64] Jianjun Li, Gehui Wang, Can Wu, Cong Cao, Yanqin Ren, Jiayuan Wang, Jin Li, Junji Cao, Limin Zeng, and Tong Zhu. Characterization of isoprenederived secondary organic aerosols at a rural site in north China plain with

implications for anthropogenic pollution effects. *Scientific Reports*, 8:535, 2018.

- [65] J. Fenger. Urban air quality. Developments in Environmental Science, 1: 1–52, December 2002.
- [66] World Health Organization. Ambient (outdoor) air quality and health, 2014. URL http://www.who.int/mediacentre/factsheets/fs313/en/. WHO fact sheet No. 313.
- [67] A. G. Barnett. It's safe to say there is no safe level of air pollution. Australian and New Zealand Journal of Public Heath, 38:407–408, October 2014.
- [68] J. Schwartz and D. W. Dockery. Increased mortality in Philadelphia associated with daily air pollution concentrations. *American review of respiratory* disease, 145(3):600–604, March 1992.
- [69] Xiping Xu, Jun Gao, and Yude Chen. Air pollution and daily mortality in residential areas of Beijing, China. Archives of Environmental Health, 49(4): 216–222, August 1993.
- [70] D. W. Dockery, C. Arden Pope, Xiping Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris, Jr., and F. E. Speizer. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*, 329:1753–1759, December 1993.
- [71] A. S. V. Shah, K. Ken Lee, D. A. McAllister, A. Hunter, H. Nair, W. Whiteley, J. P. Langrish, D. E. Newby, and N. L. Mills. Short term exposure to air pollution and stroke: systematic review and meta-analysis. *British Medical Journal*, 350:h1295, February 2015. URL http://www.bmj.com/content/ bmj/350/bmj.h1295.full.pdf.

- [72] H. Mustafi, P. Jabre, C. Caussin, M. H. Murad, S. Escolano, M. Tafflet, M. Prier, E. Marijon, D. Vernerey, J. Empana, and X. Jouven. Main air pollutants and myocardial infarction. *Journal of the American Medical Association*, 713, February 2012. URL http://jama.jamanetwork.com/article. aspx?articleid=1104975.
- [73] Jiming Hao, Dongquan He, Ye Wu, Lixn Fu, and Kebin He. A study of the emission and concentration distribution of vehicular pollutants in the urban area of Beijing. *Atmospheric Environment*, 3:435–465, February 2000.
- [74] W. B. Johnson, F. L. Ludwig, W. F. Dabberdt, and R. J. Allen. An urban diffusion simulation model for carbon monoxide. *Journal of the Air Pollution Control Association*, 23:490–498, June 1973.
- [75] M. J. Kleeman and G. R. Cass. Source contributions to the size and composition distribution of urban particulate air pollution. *Atmospheric Environment*, 32:2803–2816, August 1998.
- [76] J. P. Lawrence, J. S. Anand, J. D. Vande Hey, J. White, R. R. Leigh, P. S. Monks, and R. J. Leigh. High-resolution measurements from the airborne atmospheric nitrogen dioxide imager (ANDI). *Atmospheric Measurement Techniques*, 8(11):4735–4754, November 2015.
- [77] R. Berkowicz. A simple model for urban background pollution. Environmental monitoring and assessment, 65(1-2):259–267, November 2000.
- [78] J. Gulliver and D. Briggs. Stems-air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. Science of the Total Environment, 409:2419–2429, 2011.
- [79] D. G. Streets, J. S. Fu, C. J. Jang, Jiming Hao, Kebin He, Xiaoyan Tang, Yuanhang Zhang, Zifa Wang, Zuopan Li, Qiang Zhang, Litao Wang, Binyu Wang, and Carolyne Yu. Air quality during the 2008 Beijing Olympic games. *Atmospheric Environment*, 41:480–492, January 2007.

- [80] Healthier Air for Leicester: Leicesters Air Quality Action Plan (2015-2025). Leicester City Council, November 2015. URL https://www.leicester. gov.uk/media/180653/air-quality-action-plan.pdf.
- [81] E. J. S. Ferranti, A. R. MacKenzie, K. Ashworth, and C. N. Hewitt. First steps in urban air quality for built environment practitioners. Trees and Design Action Group, 2017. URL http://epapers.bham.ac.uk/3069.
- [82] Tingzhen Chong Ming, Peng, Tingrui Gong, and Zheng-Pollution Li. tong dispersion inbuilt environment, chapter 2. URL http://www.intechopen.com/books/ Elsevier, 2011. chemistry-emission-control-radioactive-pollution-and-indoor-air-quality/ air-quality-in-rural-areas.
- [83] K.V. Abhijith, P. Kumar, J. Gallagher, A. McNabola, and B. Broderick S. Di Sabatino B. Pulvirenti R. Baldauf, F. Pilla. Air pollution abatement performances of green infrastructure in open road and built-up street canyon environments - a review. *Atmospheric Environment*, 162:71–86, 2017.
- [84] A. P. R. Jeanjean, R. Buccolieri, J. Eddy, P. S. Monks, and R. Leigh. Air quality affected by trees in real street canyons: The case of Marylebone neighbourhood in central London. Urban Forestry and Urban Greening, 22: 41–53, March 2017.
- [85] Y B Zeldovich. The oxidation of nitrogen in combustion explosions. Acta Physicochimica U.S.S.R., 21:577–628, 1954.
- [86] M. Sillman. The relation between ozone, NO and hydrocarbons in urban and polluted rural environments. *Atmospheric Environment*, 33:1821–1845, 1999.
- [87] International Agency for Research on Cancer. Re-evaluation of some organic chemicals, hydrazene and hydrogen peroxide. In IARC Monograph on the Evaluation of Carcinogenic Risks to Humans, chapter 2. IARC, 1999.

- [88] B. W. LaFranchi, G. M. Wolfe, J. A. Thornton, S. A. Harrold, E. C. Browne, K. E. Min, P. J. Wooldridge, J. B. Gilman, W. C. Kuster, P. D. Goldan, J. A. de Gouw, M. McKay, A. H. Goldstein, X. Ren, J. Mao, and R. C. Cohen. Closing the peroxy acetyl nitrate budget: observations of acylperoxy nitrates (PAN, PPN, and MPAN) during BEARPEX 2007. Atmospheric Chemistry and Physics, 9:7623–7641, October 2009. URL http: //www.atmos-chem-phys.net/9/7623/2009/acp-9-7623-2009.pdf.
- [89] L. Clapp and M. Jenkin. Analysis of the relationship between ambient levels of O3, NO2 and NO as a function of NOx in the UK. Atmospheric Environment, 35:6391–6405, 2001.
- [90] I. Heimann, V. B. Bright, M. W. McLeod, M. I. Mead, O. A. M. Popoola, G. B. Stewart, and R. L. Jones. Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors. *Atmospheric Environment*, 113:10–19, 2013.
- [91] G. B. Stewart P. Landshoff M. Calleja M. Hayes J. J. Baldovi M. W. McLeod T. F. Hodgson J. Dicks M. I. Mead, O. A. M. Popoola. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment*, 70:186–203, 2013.
- [92] Tim Wether. A simple spectroscope based on a prism, 2007. URL https://en.wikipedia.org/wiki/Optical\_spectrometer#/media/ File:Simple\_spectroscope.jpg.
- [93] Hans Edner, Pr Ragnarson, Stefan Spnnare, and Sune Svanberg. Differential optical absorption spectroscopy (DOAS) system for urban atmospheric pollution monitoring. *Applied Optics*, 32:327–333, 1993.
- [94] U. Platt and D. Perner. Measurements of atmospheric trace gases by long path differential uv/visible absorption spectroscopy. Optical and Laser remote sensing, 39:95–105, 1983.

- [95] G. Lvblad, F. Palmgren, R van Aalst, I. Allegrini, R. Bauman, A. Derouane, L. Edwards, I. Fiala, A. Hauer, M. Hawkins, T. Lahtinen, R. van Leeuwen, M. Nilsson, V. Pohjola, E. De Saeger, and F. De Santis. Position paper on air quality: Nitrogen dioxide. Technical report, European Commission, November 1999. URL https://ec.europa.eu/environment/archives/air/pdf/ pp\_no2.pdf.
- [96] J. Lawrence. Cityscan HSI-DOAS webpage, July 2016. URL http://www2. le.ac.uk/departments/physics/research/eos/james-lawrence.
- [97] R. Graves. Can a hemispherical scanning imaging differential optical absorption spectroscopy technique be used to improve understanding of the emissions, dynamics and chemistry of nitrogen dioxide in the urban boundary layer? PhD thesis, University of Leicester, March 2015.
- [98] K. F. Boersma, D. J. Jacob, M. Trainic, Y. Rudich, I. DeSmedt, R. Dirksen, and H. J. Eskes. Validation of urban NO2 concentrations and their diurnal and seasonal variations observed from the SCIAMACHY and OMI sensors using in-situ surface measurements in israeli cities. *Atmos. Chem. Phys.*, 9: 3867–3879, 2009. URL https://doi.org/10.5194/acp-9-3867-2009.
- [99] C. E. Junge. Global ozone budget and exchange between stratosphere and troposphere. *Tellus*, 14:363–377, November 1962.
- [100] J. P. Burrows, M. Weber, M. Buchwitz, V. Rozanov, A. Ladsttter-Weienmayer, A. Richter, R. DeBeek, R. Hoogen, K. Bramstedt, K. Eichmann, M. Eisinger, and D. Perner. The global ozone monitoring experiment (GOME): Mission concept and first scientific results. *Journal of Atmospheric Science*, 56:151–175, 1999.

- [101] H. Bovensmann, J. P. Burrows, M. Buchwitz, J. Frerick, S. Nol, V. V. Rozanov, K. V. Chance, and A. P. H. Goede. SCIAMACHY: Mission objectives and measurement modes. *Journal of the Atmospheric Sciences*, 56 (2):127–150, January 1999.
- [102] J. Callies, E. Corpaccioli, M. Eisinger, A. Hahne, and A. Lefebvre. GOME2: Metops second-generation sensor for operational ozone monitoring. *ESA Bulletin*, 102:28–36, May 2000.
- [103] AVIRIS, 2018. URL https://aviris.jpl.nasa.gov/aviris/index.html. Homepage of Airborne Visible/Infrared Imaging Spectrometer.
- [104] FAAM, 2018. URL https://www.faam.ac.uk/what-is-faam/. Homepage of the Facility for Airbourne Atmospheric Measurements.
- [105] G. J. Wendel, D. H. Stedman, and C. A. Cantrell. Luminol-based nitrogen dioxide detector. Analytical Chemistry, 55:937–940, 1983.
- [106] P. Mikuka and Z. Veea. Determination of nitrogen dioxide with a chemiluminescent aerosol detector. Analytical Chemistry, 64:2187–2191, 1992.
- [107] T. J. Kelly, C. W. Spicer, and G. F. Ward. An assessment of the luminol chemiluminescence technique for measurement of NO2 in ambient air. *Atmospheric Environment*, 24A:2397–2403, 1990.
- [108] J. N. Bates. Nitric oxide measurement by chemiluminescence detection. *Neuroprotocols*, 1(2):141–149, 1992.
- [109] Chemiluminescence measurmenet of NO/NOx in gas analysers, 2018. URL www.k2bw.com/chemiluminescence.htm.
- [110] G. Villena, I. Bejan, R. Kurtenbach, P. Wiesen, and J. Kleffman. Interferences of commercial NO2 instruments in the urban atmosphere and in a smog chamber. *Atmos. Mes. Tech.*, 5:149–159, 2012.

- [111] P. L. Kebabian, E. C. Wood, S. C. Herndon, and A. Freedman. A practical alternative to chemiluminescence-based detection of nitrogen dioxide: Cavity attenuated phase shift spectroscopy. *Environ. Sci. Technol.*, 42:6040–6045, 2008.
- [112] Air Quality Expert Group. Measurement methods and uk monitoring networks for NO2. In Nitrogen Dioxide in the UK, chapter 4. Department for the Environment, Food and Rural Affairs, 2014. URL https: //uk-air.defra.gov.uk/library/aqeg/publications.
- [113] P. L. Kebabian, W. A. Robinson, and A. Freedman. Optical extinction monitor using CW cavity enhanced detection. *Review of Scientific Instruments*, 78:063102, 2007.
- [114] Serinus 60 Spec Sheet. Ecotech, 2018. URL https://www.ecotech.com/wp-content/uploads/2018/03/ ECOTECH-Serinus-60-NO2-Analyser-20180315.pdf.
- [115] A. Freedman. CAPS NO2 Monitor. Aerodyne Research, 2018. URL http: //www.aerodyne.com/products/caps-no2-monitor.
- [116] B. Stuart. Infrared spectroscopy. In Kirk-Othmer Encyclopedia of Chemical Technology, pages 1–18. American Cancer Society, September 2015. ISBN 9780471238966.
- [117] Introduction to FTIR spectroscopy. Newport, August 2016. URL https: //www.newport.com/n/introduction-to-ftir-spectroscopy.
- [118] D. S. Venables, T. Gherman, J. Orphal, J. C. Wenger, and A. R. Ruth. High sensitivity in situ monitoring of NO3 in an atmospheric simulation chamber using incoherent broadband cavity-enhanced absorption spectroscopy. *Environ. Sci. Tech.*, 40:6758–6763, 2006.

- [119] J. M. Langridge, S. M. Ball, A. J. L. Shillings, and R. L. Jones. A broadband absorption spectrometer using light emitting diodes for ultrasensitive, in situ trace gas detection. *Review of Scientific Instruments*, page 123110, 2008.
- [120] Noora Sh. Oraha Qas Nouna. Condense-phase BBCEAS measurements at UV wavelength. World Scientific News, 35:17–29, 2016.
- [121] I. Ventrillard-Courtillot, E. Sciamma OBrien, S. Kassi, G. Mjean, and D. Romanini. Incoherent broad-band cavity-enhanced absorption spectroscopy for simultaneous trace measurements of NO2 and NO3 with a LED source. Applied physics B, 79(12):123110, 2008.
- [122] Sophie Dixneuf professional webpage, July 2016. URL http://www. physics.ucc.ie/lsg/sophie.html.
- [123] MICS-5914 Datasheet. SGX Sensortech, March 2016. URL http://www.sgxsensortech.com/content/uploads/2014/07/1108\_ Datasheet-MiCS-5914.pdf.
- [124] G. F. Fine, L. M. Cavanagh, A. Afonja, and R. Binions. Metal oxide semiconductor gas sensors in environmental monitoring. *Sensors*, 10:5469–5502, 2010.
- [125] D. E. Williams, J. Salmond, Y. Yu Fai, A. Jin, B. Wright, J. Wilson, G. S. Henshaw, D. B. Wells, G. Ding, J. Wagner, and G. Laing. Development of low-cost ozone and nitrogen dioxide measurement instruments suitable for use in an air quality monitoring network. 2009 IEEE Sensors, pages 1099–1104, 2009.
- [126] I. Elmi, S. Zampolli, E. Cozzani, F. Mancarella, and G. C. Cardinali. Development of ultra-low-power consumption MOX sensors with ppb-level VOC detection capabilities for emerging applications. *Sensor. Actuat. B-Chem.*, 135:342–351, 2008.

- [127] S. Zampolli, I. Elmi, F. Mancarella, P. Betti, E. Dalcanale, G. C. Cardinali, and M. Severi. Real-time monitoring of sub-ppb concentrations of aromatic volatiles with a MEMS-enabled miniaturized gas-chromatograph. *Sensor. Actuat. B-Chem.*, 141:322–328, 2009.
- [128] D. E. Williams, S. R. Aliwell, K. F. E. Pratt, D. J. Caruana, R. L. Jones, R. A. Cox, G. M. Hansford, and J. Halsall. Modelling the response of a tungsten oxide semiconductor as a gas sensor for the measurement of ozone. *Meas. Sci. Technol.*, 13:923, 2002.
- [129] S. Bicelli, A. Depari, G. Faglia, A. Flammini, A. Fort, M. Mugnaini, A. Ponzoni, V. Vignoli, and S. Rocchi. Model and experimental characterization of the dynamic behavior of low-power carbon monoxide MOX sensors operated with pulsed temperature profiles. *Meas. Sci. Technol.*, 58:1324–1332, 2009.
- [130] A. Burresi, A. Fort, S. Rocchi, B. Serrano, N. Ulivieri, and V. Vignoli. Dynamic CO recognition in presence of interfering gases by using one MOX sensor and a selected temperature profile. *Sensor. Actuat. B-Chem.*, 106: 40–43, 2005.
- [131] Z. Lei, T. Feng-Chun, P. Xiong-Wei, and Y. Xin. A rapid discreteness correction scheme for reproducibility enhancement among a batch of MOS gas sensors. *Sensor. Actuat. A-Phys.*, 205:170–176, 2014.
- [132] A. D. Wilson. Review of electronic-nose technologies and algorithms to detect hazardous chemicals in the environment. *Proceedia Technology*, 1:453– 463, 2012.
- [133] J. E. Thompson. Crowd-sourced air quality studies: A review of the literature and portable sensors. Trends in Environmental Analytical Chemistry, 11:23–34, 2016.

- [134] M. Vezzoli, A. Ponzoni, M. Pardo, M. Falasconi, G. Faglia, and G. Sberveglieri. Exploratory data analysis for industrial safety application. *Sensor. Actuat. B-Chem.*, 131:100–109, 2008.
- [135] W. J. Buttner, M. B. Post, R. Burgess, and C. Rivkin. An overview of hydrogen safety sensors and requirements. *Int. J. Hydrogen Energ.*, 36: 2462–2470, 2011.
- [136] Boon-Brett, L.; Bousek, J.; Moretto, and P. Reliability of commercially available hydrogen sensors for detection of hydrogen at critical concentrations: part ii–selected sensor test results. Int. J. Hydrogen Energ., 34:562– 571, 2009.
- [137] How electrochemical gas sensors work. Alphasense, August 2016. URL http://www.alphasense.com/WEB1213/wp-content/uploads/2013/ 07/AAN\_104.pdf.
- [138] R. Baron and J. Saffell. Amperometric gas sensors as a low cost emerging technology platform for air quality monitoring applications: A review. ACS Sens., 2:1553–1566, 2017.
- [139] Electrochemical Sensors Application Note 2: Design of Electronics for Electrochemical Gas Sensors. SGX Sensortech, July 2016. URL https://www.sgxsensortech.com/content/uploads/2014/08/A1A-EC\_ SENSORS\_AN2-Design-of-Electronics-for-EC-Sensors-V4.pdf.
- [140] E. D. Palmes, A. F. Gunnison, J. DiMattio, and C. Tomczyk. Personal sampler for nitrogen dioxide. American Industrial Hygene Association, 37: 570–577, 1976.
- [141] P. R. Hargreaves, A. Leidi, H. J. Grubb, M. T. Howe, and M. A. Mugglestone. Local and seasonal variations in atmospheric nitrogen dioxide levels at Rothamsted, UK, and relationships with meteorological conditions. *Atmospheric Environment*, 34:843–853, 2000.

- [142] J. Targa and A. Loader. Diffusion tubes for ambient NO2 monitoring: Practical guidance. Technical report, Department for the Environment, Food and Rural Affairs, February 2008.
- [143] G. W. Campbell, J. R. Stedman, and K. Stevenson. A survey of nitrogen dioxide concentrations in the United Kingdom using diffusion tubes, July -December 1991. Atmospheric Environment, 28:477–486, February 1994.
- [144] Serinus 40 oxides of nitrogen analyser. Ecotech, July 2016. URL http://www.ecotech.com/wp-content/uploads/2016/03/ Serinus-40-Spec-Sheet.pdf.
- [145] E. Flores, J. Viallon, P. Moussay, and R. I. Wielgosz. Accurate Fourier transform infrared (FT-IR) spectroscopy measurements of nitrogen dioxide (NO2) and nitric acid (HNO3) calibrated with synthetic spectra. *Applied Spectroscopy*, 67:1171–1178, 2013.
- [146] D. Balis, M. Kroon, M. E. Koukouli, E. J. Brinksma, G. Labow, J. P. Veefkind, and R. D. McPeters. Validation of ozone monitoring instrument total ozone column measurements using Brewer and Dobson spectrophotometer groundbased observations. *Journal of Geophysical Research Atmospheres*, 112(D24), December 2007.
- [147] M. Torras and J. K. Boyce. Income, inequality, and pollution: a reassessment of the environmental Kuznets curve. *Ecological Economics*, 25:147–160, 1998.
- [148] Leicester University AURN station (UKA00573) site information, 2018. URL https://uk-air.defra.gov.uk/networks/site-info?site\_id= LECU.
- [149] Air quality guidelines for Europe. World Health Organization, 2 edition, 2000. URL http://wedocs.unep.org/bitstream/handle/20.500.11822/ 8681/Air\_quality\_guidelines.pdf.

- [150] Criteria for EUROAIRNET The EEA air quality and information network. European Environment Agency, 1999. URL https://www.eea.europa.eu/ publications/TEC12/file.
- [151] R. Williams, R. Long, M. Beaver, A. Kaufman, F. Zeiger, M. Heimbinder, B. R. Acharya, B. A. Grinwald, K. A. Kupcho, S. E. Robinson, O. Zaouak, B. Aubert, M. Hannigan, R. Piedrahita, N. Masson, B. Moran, M. Rook, P. Heppner, C. Cogar, N. Nikzad, and W. G. Griswold. Sensor evaluation report. Technical Report EPA 600/R-14/143, United States Environmental Protection Agency, May 2014.
- [152] A. C. Rai and P. Kumar. Report: Summary of air quality sensors and recommendations for application. Technical Report 0.4 WP1, European Union, February 2017.
- [153] AirBeam Technical Specification. Habitatmap, August 2016. URL http: //www.takingspace.org/aircasting/airbeam/.
- [154] Altitude. Kickstarter for Sensely, August 2016. URL https://www.kickstarter.com/projects/sensly/ sensly-a-smart-portable-pollution-sensor-for-your/description.
- [155] SensorCon, August 2016. URL http://sensorcon.com. Product website.
- [156] SensorCon company description. Grand St., August 2016. URL http:// grandst.com/u/sensorcon.
- [157] Rodrigo Narciso. Keep track of your gasses with CH4, June 2018. URL https://www.kickstarter.com/projects/963861855/ keep-track-of-your-gases-with-ch4/description. Product website.
- [158] Adafruit. Adafruit SGP30 air quality sensor breakout, June 2018. URL https://learn.adafruit.com/ adafruit-sgp30-gas-tvoc-eco2-mox-sensor/. Product website.

- [159] Air Quality System for Makers. Spec-Sensors, June 2018. URL https: //www.spec-sensors.com/air-quality-system-for-makers/.
- [160] Plotly. Arduino + air quality sensors, June 2018. URL https://plot.ly/ arduino/air-quality-tutorial/.
- [161] Make: magazine. Arduino projects, June 2018. URL https://makezine. com/category/technology/arduino/.
- [162] AQMesh. AQMesh, May 2018. URL http://www.aqmesh.com. Product website.
- [163] G. R. McKercher, J. A. Salmond, and J. K. Vanos. Characteristics and applications of small, portable gaseous air pollution monitors. *Environmental Pollution*, 223:102–110, 2017.
- [164] Aeroqual. SM50 ozone sensor circuit, May 2018. URL https://www. aeroqual.com/product/sm50-ozone-sensor-circuit. Product website.
- [165] Portable and Fixed Monitor Specifications. Aeroqual, May 2018. URL https://d2pwrbx99jwry6.cloudfront.net/wp-content/uploads/ Aeroqual-Portable-Fixed-Sensor-Specifications.pdf.
- [166] Aeroqual. Aeroqual Series 500, May 2018. URL https://www.aeroqual. com/product/series-500-portable-air-pollution-monitor. Product website.
- [167] W. Jiao, G. Hagler, R. Williams, R. Sharpe, R. Brown, D. Garver, R. Judge, M. Caudill, J. Rickard, M. Davis, L. Weinstock, S. Zimmer-Dauphinee, and K. Buckley. Community air sensor network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmos, Meas. Tech*, 9:5281–5292, 2016.

- [168] S. Moltchanov, I. Levy, Y. Etzion, U. Lerner, D. M. Broday, and B. Fishbain. On the feasibility of measuring urban air pollution by wireless distributed sensor networks. *Science of the Total Environment*, 502:537–547, 2015.
- [169] Airbase canarIT, 2018. URL http://laceen.org/laceen/airbase. Website of CanarIT pollution monitor.
- [170] C. Lin, J. Gillespie, M. D. Schuder, W. Duberstein, I. J. Beverland, and M. R. Heal. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. *Atmospheric Environment*, 100:111–116, 2015.
- [171] MICS OZ-47 Datasheet. SGX Sensortech, 2 edition, 2008. URL https: //gitlab.ethz.ch/tec/public/opensense/wikis/files/micsoz47.pdf.
- [172] MICS 2610 Datasheet. SGX Sensortech, a1a revision 2 edition, 2008.
- [173] MICS 2614 Datasheet. SGX Sensortech, 1087 revision 5 edition, 2016.
- [174] Proactive E2V620mln all-Investor. technologies agrees cash takeover offer from teledyne, 2016. URL http: //www.proactiveinvestors.co.uk/companies/news/170331/ e2v-technologies-agrees-620mln-all-cash-takeover-offer-from-teledyne-17033 html.
- [175] C. Borrego, A. M. Costa, J. Ginja, M. Amorim, M. Coutinho, K. Karatzas, Th. Sioumis, N. Katsifarakis, K. Konstantinidis, S. De Vito, E. Esposito, P. Smith, N. Andre, P.Gerard, L.A. Francis, N. Castell, P. Schneider, M. Viana, M. C. Minguillon, W. Reimringer, R. P. Otjes, O. von Sicard, R. Pohle, B. ElenD. Suriano, V. Pfister, M. Prato, S. Dipinto, and M. Penza. Assessment of air quality microsensors versus reference methods: The Eu-NetAir joint exercise. Atmospheric Environment, 147:246–263, 2016.

- [176] European network on new sensing technologies for air-pollution control and environmental sustainability: EuNetAir, 2018. URL http://www. eunetair.it.
- [177] M. Mueller, J. Meyer, and C. Hueglin. Design of an ozone and nitrogen dioxide sensor unit and its long-term operation within a sensor network in the city of Zurich. Atmos. Meas. Tech, 10:3783–3799, 2017.
- [178] Gesundheits und Umweltdepartement. Luftbilanz 2016 Resultate der Flchendeckenden Messkampagne. Umwelt und Gesundheitsschutz Stadt Zurich, 2016. URL https://www.stadt-zuerich.ch/gud/de/index/ umwelt\_energie/luftqualitaet/publikationen/luftbilanz-2016. html.
- [179] Unitec S.R.L. SENS-IT, May 2018. URL http://www.unitec-srl.com/ site/products/sens-it/. Product website.
- [180] Unitec. SENS-IT Datasheet, revision 1 edition, June 2018. URL http://www.unitec-srl.com/site/wp-content/uploads/2014/11/ SENS-IT-Datasheet\_06\_18.pdf.
- [181] L. Spinelle, M. Gerboles, M. Aleixandre, and F. Bonavitacola. Evaluation of metal oxide sensors for the monitoring of O3 in ambient air at ppb level. *Chemical Engineering Transactions*, 54:319, 2016.
- [182] Ozone: Alphasense 4-Electrode gas sensors operate using proven fuel cell technology. Alphasense, 2018. URL http://www.alphasense.com/index. php/products/ozone-2/.
- [183] OX-A431 Datasheet. Alphasense, 2017. URL http://www.alphasense. com/WEB1213/wp-content/uploads/2017/07/0X-B431.pdf.
- [184] OX-A431 Datasheet. Alphasense, 2017. URL http://www.alphasense. com/WEB1213/wp-content/uploads/2017/03/OX-A431.pdf.

- [185] N. Castell, F. R. Dauge, P. Schneider, M. Vogt, U. Lerner, B. Fishbain, D. Broday, and A. Bartonova. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International*, 99:293–302, 2017.
- [186] O3-3E1F sensor datasheet. Sensoric, 02/2011 edition, 2018. URL https: //www.citytech.com/PDF-Datasheets/o33e1f.pdf.
- [187] A3OZ sensor datasheet. Sensoric, 3 edition, 2018. URL https://www. citytech.com/PDF-Datasheets/a3oz.pdf.
- [188] L. Spinelle, M. Gerboles, M. G. Villani, M. Aleixandre, and F. Bonavitacola. Field calibration of a cluster of low-cost available sensors for air quality monitoring. part A: Ozone and nitrogen dioxide. Sensors and Actuators B: Chemical, 215:249–257, 2015.
- [189] MICS 2710 Datasheet. SGX Sensortech, 2 edition, 2008. URL https://
  www.cdiweb.com/datasheets/e2v/mics-2710.pdf.
- [190] MICS 4514 Datasheet. SGX Sensortech, 0278 revision 16 edition, 2017. URL https://www.sgxsensortech.com/content/uploads/2014/ 08/0278\_Datasheet-MiCS-4514-rev-16.pdf.
- [191] A. Bigi, M. Mueller, S. K. Grange, G. Ghermandi, and C. Hueglin. Performance of NO, NO2 low cost sensors and three calibration approaches within a real world application. *Atmos, Meas. Tech*, 11:3717–3735, 2018.
- [192] NO2-3E50 datasheet. Sensoric, 2018. URL https://www.citytech.com/ PDF-Datasheets/no23e50.pdf.
- [193] Noboru Yamazoe, Jun Fuchigami, Masato Kishikawa, and Tetsuro Seiyama. Interactions of tin oxide surface with O2, H2O and H2. Surface Science, 86: 335–344, 1979.

- [194] E. Vitz and H. Chan. LIMSport VII. semiconductor gas sensors as GC detectors and "breathalyzers". J. Chem. Educ, 72:920, 1995.
- [195] H. Meixner and U. Lampe. Metal oxide sensors. Sensor. Actuat. B-Chem., 33:198–202, 1996.
- [196] Hong-Ming Lin, Shah-Jye Tzeng, Peng-Jan Hsiau, and Wen-Li Tsai. Electrode effects on gas sensing properties of nanocrystaline zinc oxide. Nanostructured Materials, 10:465–477, 1998.
- [197] N. Yamazoe, Y. Kurokawa, and T. Seiyama. Effects of additives on semiconductor gas sensors. *Sensors and Actuators*, 4:283–289, 1983.
- [198] M. Batzill and U. Diebold. The surface and materials science of tin oxide. Progress in Surface Science, 79:47–154, 2005.
- [199] Hyo-Joong Kim and Jong-Heun Lee. Highly sensitive and selective gas sensors using p-type oxide semiconductors: Overview. Sensors and Actuators B: Chemical, 192:607–627, 03 2014.
- [200] S. C. Naisbitt, K. F. E. Pratt, D. E. Williams, and I. P. Parkin. A microstructural model of semiconducting gas sensor response: The effects of sintering temperature on the response of chromium titanate (CTO) to carbon monoxide. *Sensor. Actuat. B-Chem.*, 114:969–977, 2006.
- [201] D. E. Williams and K. F. E. Pratt. Microstructure effects on the response of gas-sensitive resistors based on semiconducting oxides. *Sensors and Actuators B: Chemical*, 70:214–221, 2000.
- [202] P. J. D. Peterson, A. Aujla, K. H. Grant, A. G. Brundle, M. R. Thompson, J. Vande Hey, and R. J. Leigh. Practical use of metal oxide semiconductor gas sensors for measuring nitrogen dioxide and ozone in urban environments. *Sensors*, 17(7):1653, 2017.

- [203] N. Masson, R. Piedrahita, and M. Hannigan. Approach for quantification of metal oxide type semiconductor gas sensors used for ambient air quality monitoring. *Sensor. Actuat. B-Chem.*, 208:339–345, 2015.
- [204] Envirologger. Envirologger AQMesh info page, May 2018. URL https: //www.envirologger.com/aq\_mesh.
- [205] Z. Fleming. Personal conversation, 2018.
- [206] CAIRPOL CAIRSENSE miniature air quality monitoring sensors network, 2018. URL http://www.altechusa.com/ cairpol-miniature-air-quality-monitoring-sensors-network. Product series website.
- [207] M. Aleixandre L. Spinelle, M. Gerboles. Performance evaluation of amperometric sensors for the monitoring of O3 and NO2 in ambient air at ppb level. *Proceedia Engineering*, 129:480–483, 2015.
- [208] R. M. Duvall, R. W. Long, M. R. Beaver, K. G. Kronmiller, M. L. Wheeler, and J. J. Szykman. Performance evaluation and community application of low-cost sensors for ozone and nitrogen dioxide. *MDPI Sensors*, 16:1698, 2016.
- [209] A. Cattaneo, M. Taronna, G. Garramone, C. Peruzzo, C. Schlitt, D. Consonni, and D. M. Cavallo. Comparison between personal and individual exposure to urban air pollutants. *Aerosol Science and Technology*, 44:370– 379, 2010.
- [210] K. Elampari, T. Chithambarathanu, and R. Krishna Sharma. Examining the variations of ground level ozone and nitrogen dioxide in a rural area influenced by brick kiln industries. *Indian Journal of Science and Technology*, 3:900–903, 2010.

- [211] Development of low-cost ozone and nitrogen dioxide measurement instruments suitable for use in an air quality monitoring network, Christchurch, New Zealand, October 2009. ECS Transactions, IEEE.
- [212] Brandt Scott. Utilizing portable air quality monitors to assess the patterns of ozone along the northern front range of colorado. Soars, Union College Schenectady, 2014.
- [213] Mobile Sensing Technology. UPOD, 2018. URL http:// mobilesensingtechnology.com/. Product website.
- [214] Perkin-Elmer. Perkin-Elmer Elm, April 2015. URL https://web.archive. org/web/20150423171244/http://elm.perkinelmer.com/? Product website.
- [215] South Coast AQMD. Perkin-Elmer Elm, June 2016. URL http://www. aqmd.gov/aq-spec/product/perkin-elmer---elm. Air quality sensor performance evaluation sensor report.
- [216] N. Chantrell. Air Quality Egg community air quality monitoring. Product review, February 2013. URL https://nathan.chantrell.net/20130210/ air-quality-egg-community-air-quality-monitoring/.
- [217] James Bruggers. Pollution micro-monitors less useful than hoped. Technical report, Louisville Courier-Journal, September 2014. URL http: //www.courier-journal.com/story/tech/science/environment/2014/ 09/29/pollution-micro-monitors-less-useful-hoped/16420565/.
- [218] Air Quality Egg version 2, July 2016. URL http://shop.wickeddevice. com/product/air-quality-egg-v2-no2-co/. Product website.
- [219] A594 roadside AURN station (UKA00609) site information, 2018. URL https://uk-air.defra.gov.uk/networks/site-info?site\_id=LEIR.

- [220] K. H. Grant. Development and calibration of a low cost air quality monitor and its subsequent deployment to measure the decay of air quality in vehicles and buildings by the roadside. Master's thesis, University of Leicester, 2014.
- [221] R. Thalman, M. T. Baeza-Romero, S. M. Ball, E. Borrs, M. J. S. Daniels, I. C. A. Goodall, S. B. Henry, T. Karl, F. N. Keutsch, S. Kim, J. Mak, P. S. Monks, A. Muoz, J. Orlando, S. Peppe, A. R. Rickard, M. Rdenas, P. Snchez, R. Seco, L. Su, G. Tyndall, M. Vzquez, T. Vera, E. Waxman, and R. Volkamer. Instrument intercomparison of glyoxal, methyl glyoxal and NO2 under simulated atmospheric conditions. *Atmos. Meas. Tech.*, 8: 1835–1862, 2015.
- [222] A. C. Lewis, J. D. Lee, P. M. Edwards, M. D. Shaw, M. J. Evans, S. J. Moller, K. Smith, J. W. Buckley, M. Ellis, S. Gillot, and A. White. Evaluating the performance of low cost chemical sensors for air pollution research. *Faraday Discussions*, December 2015.
- [223] NumPy, 2018. URL https://numpy.org/. Homepage of Python data analysis package.
- [224] M. Hewitson. Lisa technology package data analysis. Technical report, Max Planck Society, 2018. URL https://www.elisascience.org/ltpda/.
- [225] M. Hewitson, M. Armano, M. Benedetti, J. Bogenstahl, D. Bortoluzzi, P. Bosetti, N. Brandt, A. Cavalleri, G. Ciani, I. Cristofolini, M. Cruise, K. Danzmann, I. Diepholz, R. Dolesi, J. Fauste, L. Ferraioli, D. Fertin, W. Fichter, A. Garca, C. Garca, A. Grynagier, F. Guzmn, E. Fitzsimons, G. Heinzel, D. Hollington, J. Hough, M. Hueller, D. Hoyland, O. Jennrich, B. Johlander, C. Killow, A. Lobo, D. Mance, I. Mateos, P. W. Mc-Namara, A. Monsky, D. Nicolini, D. Nicolodi, M. Nofrarias, M. Perreur-Lloyd, E. Plagnol, G. D. Racca, J. Ramos-Castro, D. Robertson, J. Sanjuan, M. O. Schulte, D. N. A. Shaul, M. Smit, L. Stagnaro, F. Steier, T. J. Sumner, N. Tateo, D. Tombolato, G. Vischer, S. Vitale, G. Wanner, H. Ward,

S. Waschke, V. Wand, P. Wass, W. J. Weber, T. Ziegler, and P. Zweifel. Data analysis for the LISA technology package. *Classical and Quantum Gravity*, 26(9):094003, April 2009.

- [226] D. W. Fahey, G. Huebler, D. D. Parrish, E. J. Williams, R. B. Norton, B. A. Ridley, H. B. Singh, S.C. Liu, and F. C. Fehsenfel. Reactive nitrogen species in the troposphere' measurements of NO, NO2, HNO3, particulate nitrate, peroxyacetyl nitrate (PAN), O3, and total reactive odd nitrogen (NOy) at Niwot Ridge, Colorado. *Journal of Geophysical Research*, 91:9781–9793, 1986.
- [227] D. M. Hawkins. The problem of overfitting. J. Chem. Inf. Comput. Sci., 44: 1–12, 2004.
- [228] M. Aleixandre, M. del Carmen Horrillo, M. Gerboles, L. Spinelle, and F. Bonavitacola. Comparison of tin dioxide ozone sensor operated in on/off switching temperature cycles and at constant temperature. *Chemical Engineering Transactions*, 54:49, 2016.
- [229] N. Zimmerman, A. A. Presto, S. P. N. Kumar, J. Gu, A. Hauryliuk, E. S. Robinson, A. L. Robinson, and R. Subramanian. A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmospheric Measurement Techniques*, 11:291–313, 2018.
- [230] L. Breiman. Random forests. Machine Learning, 45:5–32, 2001.
- [231] R. Berkowitz, F. Palmgren, O. Hertel, and E. Vignati. Using measurements of air pollution in streets for evaluation of urban air quality-meterological analysis and model calculations. *The Science of the Total Environment*, 189/190:259–265, 1996.
- [232] M. C. Somerville, S. Mukerjee, and D. L. Fox. Estimating the wind direction of maximum air pollutant concentration. *Environmetrics*, 7:231–243, 1996.

- [233] J. R. Garratt. Review: the atmospheric boundary layer. Earth Science Reviews, 37:89–134, 1994.
- [234] T. Tersvirta. Specification, estimation, and evaluation of smooth transition autoregressive models. Journal of the American Statistical Association, 89 (425):208–218, 1994.
- [235] A. P. Brundle. The development of low cost metal-oxide semiconductor sensors for the detection of nitrogen dioxide. Master's thesis, University of Leicester, 2015.
- [236] K. J Laidler. The development of the Arrhenius equation. Journal of Chemical Education, 61(6):494, 1984.
- [237] J. Cyrys, M. Eeftens, J. Heinrich, C. Ampe, A. Armengaud, R. Beelen, T. Bellander, T. Beregszaszi, M. Birk, G. Cesaroni, and M. Cirach. Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. *Atmospheric Environment*, 62: 374–390, 2012.
- [238] E. Lebret, D. Briggs, H. van Reeuwijk, P. Fischer, K. Smallbone, H. Harssema, B. Kriz, P. Gorynski, and P. Elliott. Small area variations in ambient NO2 concentrations in four european areas. *Atmospheric Envri*onment, 34:177–185, 2000.
- [239] Quality Assurance and Quality Control (QA/QC) Procedures for UK Air Quality Monitoring under 2008/50/EC and 2004/107/EC. Department for the Environment, Food and Rural Affairs, September 2013. URL https://uk-air.defra.gov.uk/assets/documents/reports/ cat13/1309300904\_All-Networks\_QAQC\_Document\_2012\_\_Issue1a.pdf.
- [240] S. E. Gryning, A. A. M. Holtslag, J. S. Irwin, and B. Sivertsen. Applied dispersion modelling based on meteorological scaling parameters. *Atmospheric environment*, 21:79–89, 1987.

- [241] I. Lagzi, R. Meszaros, G. Gelybo, and A. Leelossy. Atmospheric chemistry. Technical report, Etvs Lornd University, 2013. URL http://elte.prompt. hu/sites/default/files/tananyagok/AtmosphericChemistry/.
- [242] Ying Zhou and J. Levy. Factors influencing the spatial extent of mobile source air pollution impacts: a meta-analysis. BMC public health, 7:89, 2007.
- [243] Earthsense Ltd. Earthsense Zephyr, 2018. URL https://www.earthsense. co.uk/zephyr. Product website.