Science of the Total Environment 575 (2017) 1119–1129

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Low-cost sensors and crowd-sourced data: Observations of siting impacts on a network of air-quality instruments

Georgia Miskell [a,b,*], Jennifer Salmond [b], David E. Williams [a]

[a] MacDiarmid Institute for Advanced Materials and Nanotechnology, School of Chemical Sciences, 23 Symonds Street, University of Auckland, New Zealand
[b] School of Environment, 10 Symonds Street, University of Auckland, New Zealand

HIGHLIGHTS

- Intra-site sensor variability was small indicating low sensitivity to siting type.
- Short-term local activities were identified but did not significantly impact reporting scales.
- Crowd-sourced sites in proximity to regulatory analyzers measured similar trends.
- With quality control checks, crowd-sourced networks provided useful additional data.

GRAPHICAL ABSTRACT

ARTICLE INFO

Article history:
Received 15 August 2016
Received in revised form 14 September 2016
Accepted 22 September 2016
Available online 29 September 2016

Editor: D. Barcelo

Keywords:
Ozone
Air quality network
Siting effect
Regional variation
Local variability
Citizen science

ABSTRACT

Low-cost sensors offer the possibility of gathering high temporal and spatial resolution crowd-sourced data-sets that have the potential to revolutionize the ways in which we understand individual and population exposure to air pollution. However, one of the challenges associated with crowd-sourced data ('citizen science'), often from low-cost sensors, is that citizens may use sites strongly affected by local conditions, limiting the wider significance of the data. This paper examines results from a low-cost network measuring ground-level ozone to evaluate the impact of siting on data quality. Locations at both reference stations and at private homes or research centers were used, and thought of as a typical 'crowd-sourced' network. Two instruments were co-located at each site to determine intra-site variability and evaluated by standard performance statistics and local-scale activity logs. The wider application of the data for both regional Inter-site variability was evaluated to show-case the wider value and usefulness of crowd-sourced data. Analysis of intra-site variability showed little differences at most sites (<5 ppb). Large differences in intra-site variability were detected when sensors were exposed to direct sunlight (causing thermal variations within the instrument) and proximity to large emission sources. Short-term local activities, such as lawn-mowing, were identifiable in the data, but had minimal impact on standard reporting time-scales, and so did not pose as being significant limitations or errors. Inter-site evaluation demonstrated that dense networks of low-cost sensors can add value to existing networks, with minimal impact on the

* Corresponding author at: MacDiarmid Institute for Advanced Materials and Nanotechnology, School of Chemical Sciences, 23 Symonds Street, University of Auckland, New Zealand.
E-mail address: gmis011@aucklanduni.ac.nz (G. Miskell).

http://dx.doi.org/10.1016/j.scitotenv.2016.09.177

overall data-set quality. Sensors located in crowd-sourced locations nearby to regulatory analyzers were able to capture similar trends and concentrations, supporting their ability to report on wider conditions. Thus crowd-sourced approaches to monitoring (with suitable calibration and data quality control checks) may be an appropriate method for increasing the temporal and spatial resolution of air quality networks.

1. Introduction

Recent air quality research has focused on different approaches in using data from low-cost instruments to supplement data provided by official regulatory bodies (Snyder et al., 2013). Whilst low-cost instruments have the potential to make a significant contribution to our understanding of the temporal and spatial variation of air pollutant concentrations in urban areas, concerns over their accuracy and precision have limited their widespread use (Ottinger, 2010; Snyder et al., 2013; Tregidgo et al., 2013). However, recent innovations in techniques to detect sensor error and improve accuracy (e.g. Alavi-Shoshtari et al., 2013; Miskell et al., 2016) are proving increasingly successful, and attention is now moving away from assessing their reliability towards developing best-practice guidelines for the use of this new technology (Nieuwenhuijsen et al., 2015; Xiang et al., 2016, U.S. Environmental Protection Agency, n.d).

One area that has been given little attention so far is the impact of local siting on determining the spatial and temporal representativeness of the data. If low-cost, crowd-sourced data is to be adopted in air quality research, then the impact upon measured concentrations of siting instruments on private homes or education centers needs to be understood. Traditionally, strict regulations surround the siting of regulatory monitoring locations to ensure that datasets are representative of a given area or land-use type and local-scale effects are controlled for (Ministry for the Environment, 2009; U.S. Environmental Protection Agency, 2013). For example, recommendations typically include that the instrument is not adjacent to any walls, avoidance of large trees, certain facades (e.g. wood), and chemical interferences (e.g. vehicle emissions), and above the urban canopy layer (Ministry for the Environment, 2009; Moosavi et al., 2015).

Citizen science approaches which may see instruments located on education centers or private homes or in gardens could provide complementary information to regulatory datasets about the effects of different land-use and settings in previously unmonitored locations (Brienza et al., 2015; Ho et al., 2014). However, they can be expected to violate a number of siting recommendations because of power requirements, aesthetics, and household surroundings (e.g. building material). Data from instruments at poorly selected locations (which may occur in crowd-sourced data due to the siting often being outside of the data users control) may not be representative of wider conditions due to dominant effects of extremely local conditions or events specific to that site. This has the potential to make data from these sites unsuitable for reporting from a network perspective, and any temporal or spatial averaging of the data could be misleading from air quality management perspectives. There is therefore a need to assess the impact of different types of siting and to develop quality assurance techniques to allow the citizen scientist (and those using that data) to know how to interpret, and what value to place on, the data from their instrument (Bonney et al., 2016; Ho et al., 2014; Wolters et al., 2016).

This paper examines the effect of local siting on data quality to address the overall enquiry on the usefulness of low-cost data. Data from a network of instruments (mounted on a variety of siting options, such as on regulatory stations or on walls of private houses) were analyzed for their intra- (within a site) and inter- (between sites) variability. Differences within a site were compared to their surroundings using regression and standard statistical diagnostics to ascertain whether certain factors were related to large intra-site differences. Factors with large differences could then be recommended to the citizen scientist to avoid when mounting an instrument, or to the data user in deciding whether to include the site within network analysis. Inter-site analysis examined how a crowd-sourced network can assist in developing and improving our understanding of the temporal and spatial variability of urban O_3 by using standard statistical diagnostics. Finally, differences between crowd-sourced sites to nearby reference stations were analyzed for their ability to capture the wider pattern and to give support for providing data representative of an area.

2. Materials and methods

2.1. Data

The data used here were collected from a network of low-cost instruments measuring ground-level ozone (O_3) around Auckland, New Zealand, over a twelve-month period (November 2014–November 2015) with two instruments operating per site (<2 m distance apart). The data were validated by using methods described previously, with good quality data capture for over 75% of all observations (Bart et al., 2014; Williams et al., 2013). Auckland has a subtropical oceanic climate, with humid summers and mild winters and prevailing wind direction from the Southwest (Adeeb and Shooter, 2004). O_3 is a secondary pollutant formed from the photochemical reaction of NO_X or VOCs with UV, which causes regular spatial profiles and so regional patterning can be expected from synoptic weather patterns up and downwind of urban centers or central business districts (CBD) where precursors are produced (often traffic-related) (Bart et al., 2014). O_3 concentrations in Auckland are typically low compared to other urban centers due to titration from nitrous gases along with the geographic setting (Jiang et al., 2014), with a peak of O_3 in the winter to spring months (July–October), believed to be from greater stratospheric intrusion rather than local sources (Adeeb and Shooter, 2004). High O_3 days occur at different times at different locations across Auckland, suggesting the significance of local-scale controls (Adeeb and Shooter, 2004). Auckland has three reference stations measuring O_3 (Fig. 1), with two (Musick Point, MP, and Whangaparaoa, WHA) operating only during the summer months. Therefore, our understanding of O_3 throughout the year in Auckland is determined from one site (Patumahoe, PAT).

The low-cost sensors used were Aeroqual gas-sensitive semiconducting oxide (GSS), which have been successfully used in a number of field studies (Bart et al., 2014; Deville Cavellin et al., 2016; Lin et al., 2015) and were found to have good performance when compared against other commercial low-cost instruments (SCAMD, 2015). Sensor specifications state a level of accuracy to 5 ppb, which has been used here as a benchmark threshold as a true, or real, difference between co-located measurements. Previous work (Bart et al., 2014) found a standard error of 6 ppb when devices were compared to co-located analyzer stations for over 6000 measurements, giving support for this level of precision of the device. All low-cost instruments were field linearized and adjusted data reported, with detailed information on corrections presented in the Supporting material. The methods described by Bart et al., 2014 were used to check instrument performance; the sensor assembly in the instrument was replaced when a signal or baseline drift was detected (typically every 2–5 months; median 3 months). Overall, the response was good, with high linearity and no significant differences between co-located concentrations following calibration. Site locations are illustrated in Fig. 1, and include both reference ($n = 3$)

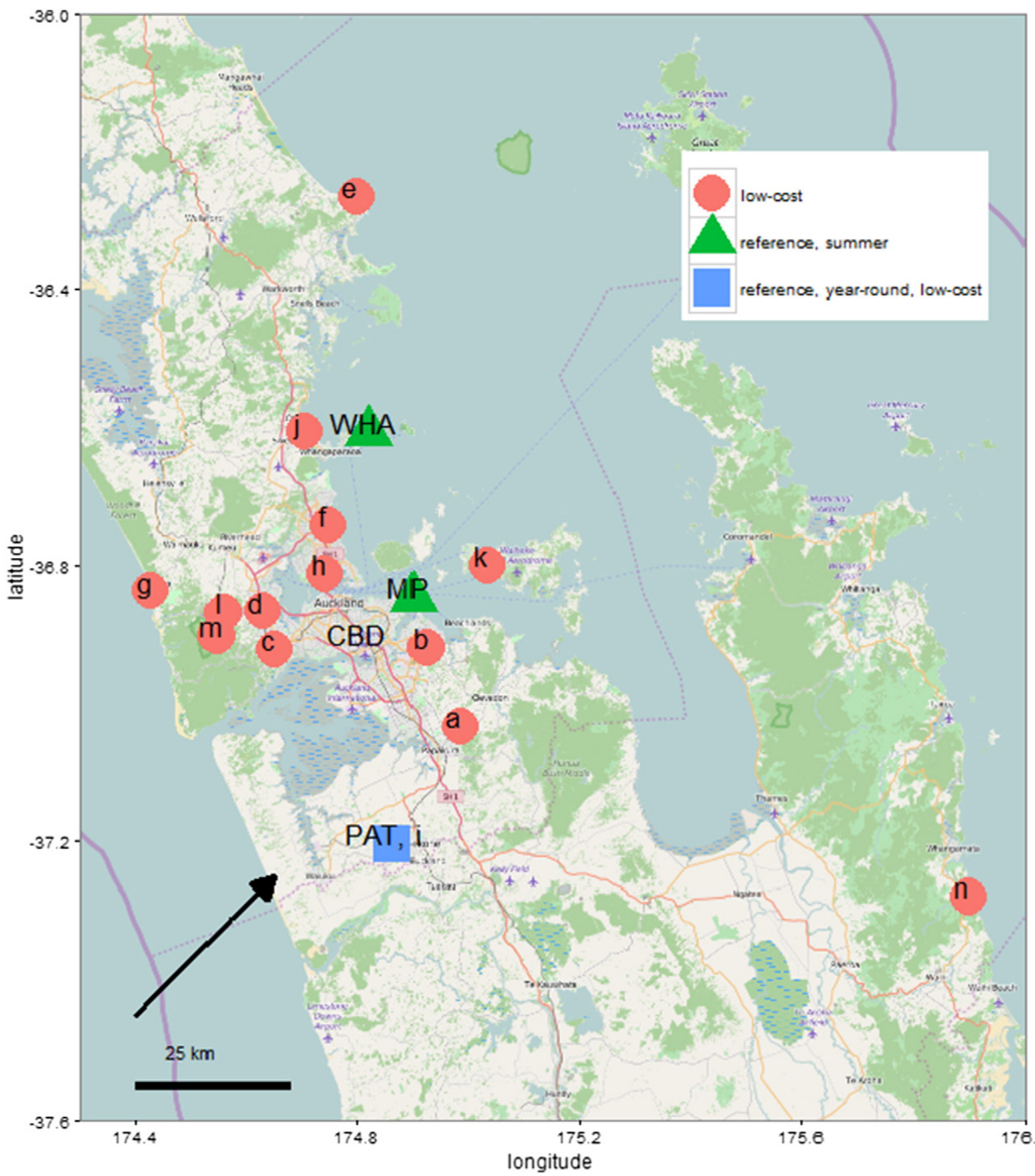

Fig. 1. Locations of monitored sites around Auckland. The arrow denotes prevailing wind direction and CBD is the urban center. Sites a – n are the low-cost sites, PAT is Patumahoe, MP is Musick Point, and WHA is Whangaparaoa.

and low-cost ($n = 14$; labelled Sites *a*–*n*) sites. Four of the low-cost sites (Sites *b*, *c*, *d*, and *i*) were co-located with reference stations (Site *i* being the only one measuring O_3), with the remainder affixed to walls or balconies in private homes or within university grounds. Periods of missing data were observed over time and were typically due to transmission or power issues, <75% daily capture, or the site not yet being established or was disestablished. Changes between co-located sensors' patterns prompted a change and recalibration of the instruments, with average length of time around 90 days. Locations where sensors were changed at <30 days or at over 90 days did not seem to be related to their site type, and so may not be responsible for drift (drift being typical in most low-cost instruments over time).

2.2. Intra-site analysis methods

2.2.1. Intra-site differences

In order to determine the effect of local conditions on the sensors, each site was classified according to a number of land-use parameters such as distance from emission sources and type of mounting (Table 1). These explanatory variables were determined based on both known O_3 sources and quantifiable observations that could help in describing immediate site surroundings. The subset of instruments co-located with stations allowed comparison between two siting types, with the co-location sites previously selected by air quality managers as ideal locations for monitoring and the other sites as a pseudo crowd-sourced

Table 1
Descriptors for each of the low-cost sites using set explanatory variables (V^x).

Site	#n (days)	Siting	V^a	V^b	V^c	V^d	V^e	V^f	V^g	V^h	V^i	V^j	V^k	V^l
a	67	Wall	1.5	S	SE	Yes	No	No	No	Grass	No	Research	Agricultural	Yes
b	347	Reference	3.5	N	SE	No	No	Yes	Yes	Grass	No	School	Residential	No
c	79	Reference	3.5	E/W	SW	No	No	Yes	No	Gravel	No	Park	Residential	No
d	294	Reference	3.5	N/S	NW	No	No	Yes	Yes	Gravel	No	School	Residential	No
e	345	Wall/Roof	2	N	NE	No	No	No	No	Deck	Yes	Research	Coastal	Yes
f	329	Wall	1.5	N	NW	No	Yes	No	No	Gravel	Yes	House	Residential	Yes
g	224	Balcony	4	E/W	NW	Yes	Yes	No	No	Deck	Yes	House	Residential	No
h	231	Wall	2.5	W	NW	Yes	Yes	Yes	No	Grass	Yes	House	Residential	Yes
i	287	Reference	3.5	W	SE	No	No	No	No	Grass	No	Research	Agricultural	No
j	336	Balcony	3	N	NW	Yes	Yes	No	No	Grass	Yes	House	Residential	Yes
k	324	Wall	1.5	E	NE	No	Yes	Yes	No	Grass	Yes	Research	Agricultural	Yes
l	334	Wall	1.5	W	SW	No	Yes	Yes	No	Gravel	No	House	Bush	Yes
m	122	Balcony	4	E	SW	No	Yes	Yes	No	Grass	No	House	Bush	Yes
n	217	Balcony	2	E	SE	No	Yes	Yes	No	Gravel	Yes	House	Coastal	Yes

[a] Height above ground (m).
[b] Direction the instrument is facing.
[c] Direction from the CBD.
[d] Shelter over instrument.
[e] Tree within 10 m.
[f] Small emission source within 10 m.
[g] High emission source within 10 m.
[h] Land coverage below the instrument.
[i] Water-body within 1000 m.
[j] Descriptor of location.
[k] Land-use designation ('Bush' is North Island NZ native forest).
[l] Proximity to wood surface.

network where locations would be outside of a managers' control (here labelled as station – controlled and crowd-sourced – uncontrolled). Co-located data were analyzed using commonly used statistics to assess performance on a number of indicators at both hourly and daily scales (mean absolute error – MAE and root mean square error – RMSE for accuracy, coefficient of determination – R^2 and spearman rank correlation – ρ for precision and relative ranking, and Cohen's d scores (Cohen, 1988) for practical significance on the size of the effect). Grouping the concentration differences into low (<5 ppb) and high bands (based on thresholds previously determined, Bart et al., 2014) further

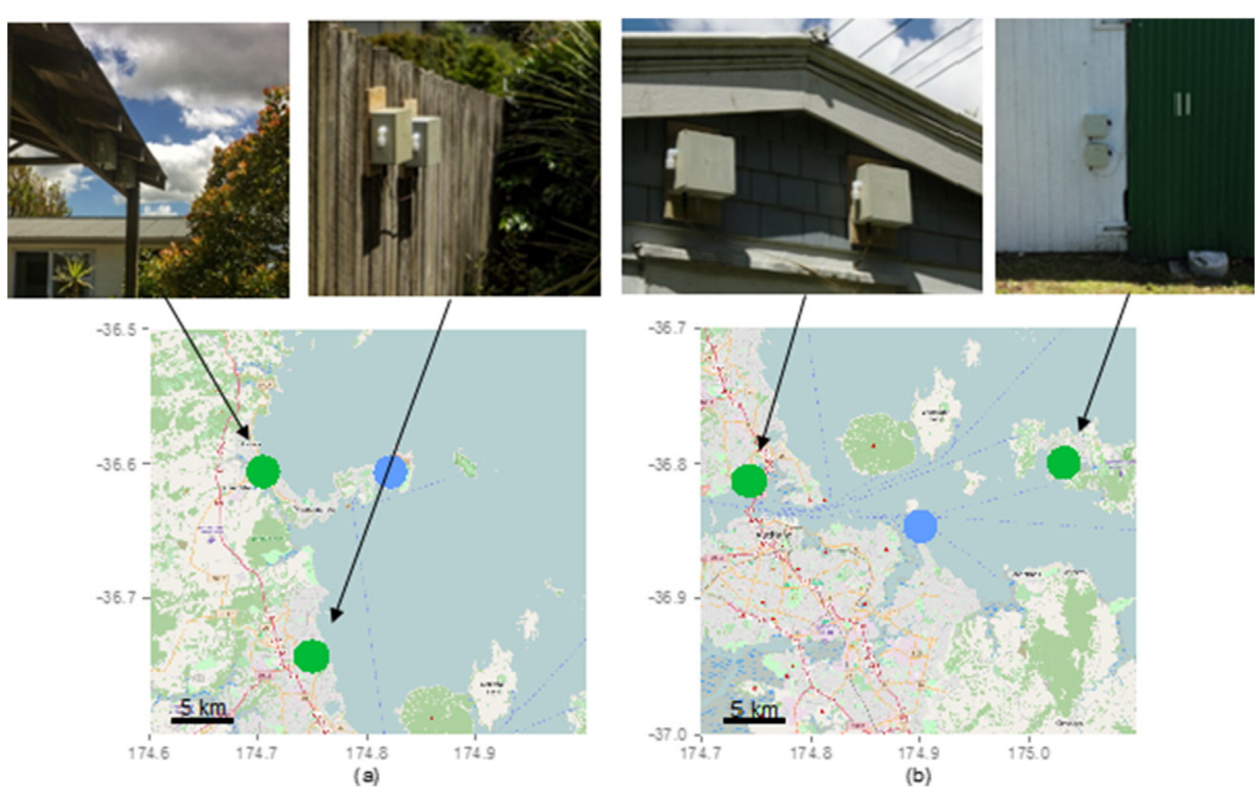

Fig. 2. Locations of the low-cost sites (green) in close proximity to a reference station (blue), and photographs of the installations. Figure (a) is for the Whangaparaoa analyzer and Sites f and j, and Figure (b) is for the Musick Point analyzer and Sites h and k.

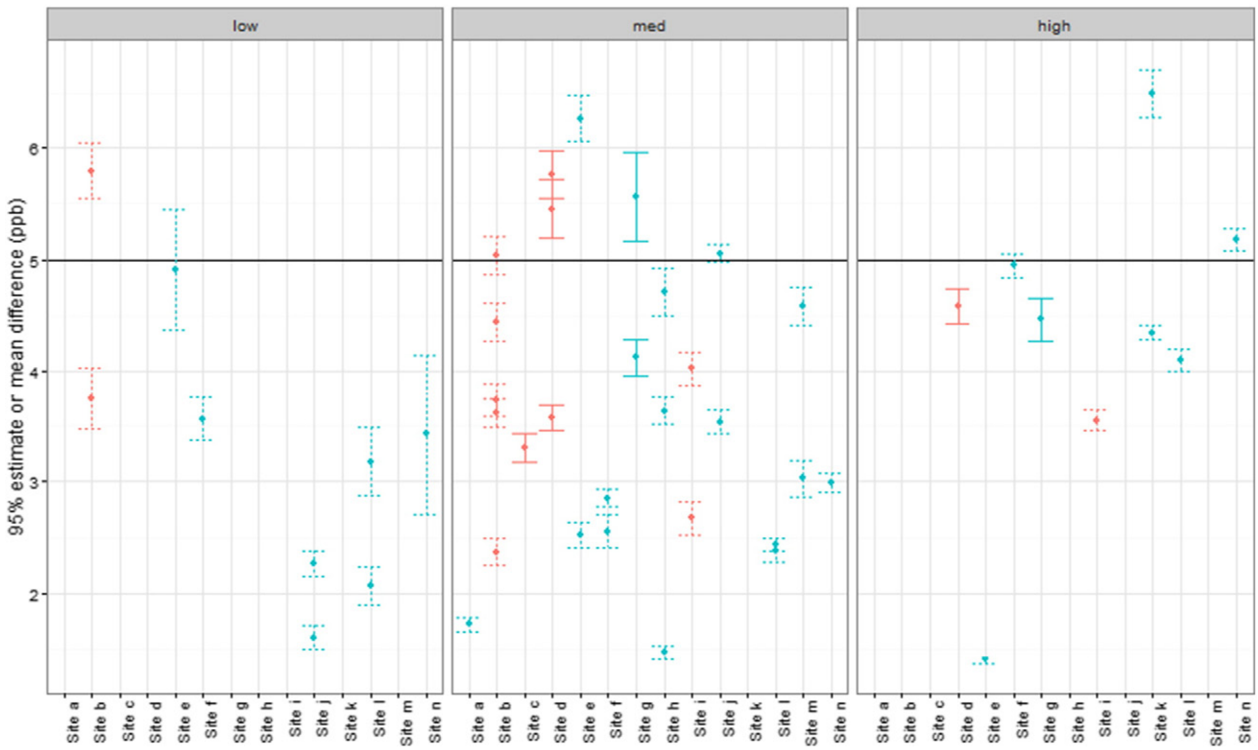

Fig. 3. 95% estimate on mean absolute difference for the 14 low-cost sites. Red bars denote where siting was controlled and blue where the siting was uncontrolled. Data were broken into bands based on measuring length of time, with low (0–30 days), med (31–90 days), and high lengths (91 + days). The threshold of 5 ppb was used for determining practical differences (Bart et al., 2014).

provided information on site surroundings and instrument functioning at real concentration levels, with variables associated with higher differences then recommended to avoid when mounting an instrument

2.2.2. Local-scale impacts

Analysis of data from local-scale activities that are known to impact short-term O_3 concentration were examined to determine if they could be distinguished from regular periods of monitoring (and therefore assist in ensuring data quality when short-term activities do occur at uncontrolled sites). Two examples were used, where one considered the impact of a short-term activity over both a spatial and temporal scale, and the other considered the impact of a longer-term seasonal activity associated with a site over a temporal scale.

The first example used co-located sensors that were placed above a grass area where lawn mowing was actively noted (two separate occasions noted at two different sites). Lawn mowing can often use gas-powered devices, which produce NO_X and therefore impact the air at a site over a small time-scale (minutes). This type of activity, although accurate for the site at that brief moment in time, is often extremely localized and may affect averaged concentrations at which data is reported due to their 'spike' impacts. Diagnostics from the sensors were analyzed to also help with identification of spikes. In particular, we used the sensor resistance baseline to check for stability of the sensor's zero over time as spikes were often associated with a type of chemical interference. In addition, we compared O_3 concentrations to the nearest site in a similar land-use setting over the same time-scale to understand the spatial impact of the activity. Impacts on the data at official reporting averages (here one-hour and rolling eight-hours) were completed by invalidating O_3 at the peak impact of the activity and comparing this to where no data were removed. This allowed for the lawn-mowing impact to be quantified to assess if such activities created significant effects at typical reporting scales.

Table 2
Intra-site regression results for the logged mean absolute difference in concentrations for each instrument/site combination and each descriptor variable from Table 1 ($n = 46$).

	Siting	V^a	V^b	V^c	V^d	V^e	V^f	V^g	V^h	V^i	V^j	V^k	V^l
Intercept	1.4**	1**	1.4**	1.3**	1.3**	1.3**	1.2**	1.2**	1.3**	1.3**	1.2**	1.4**	1.4**
Coef. A^	Balcony: Base	0.11	E: Base	NE: Base	No: Base	No: Base	No: Base	No: Base	Deck: Base	No: Base	Coast: Base	Agricultural: Base	No: Base
Coef. B^	Roof: 0.55	–	N: −0.19	NW: −0.05	Yes: −0.15	Yes: −0.07	Yes: 0.12	Yes: 0.25	Grass: −0.09	Yes: −0.01	Park: 0.01	Bush: −0.16	Yes: −0.25
Coef. C^	Reference: 0.1	–	S: −0.87*	SE: −0.07	–	–	–	–	Gravel: −0.08	–	Research: 0.03	Coast: −0.02	–
Coef. D^	Wall: −0.18	–	W: −0.34	SW: −0.23	–	–	–	–	–	–	School: 0.26	Residential: 0.05	–
p-Value	0.08	0.08	0.05	0.69	0.39	0.54	0.3	0.07	0.86	0.98	0.34	0.67	0.03
R^2	0.15	0.07	0.23	0.03	0.02	0.01	0.02	0.08	0.01	0	0.08	0.04	0.1
χ^2^^	0.21^^^	0.76	0.33	0.43	1	0.66	0.87	0.24	0.81	1	0.4	0.51	0.46

^ Coef. A–D represent the different variable factors within the regression, with each cell identifying the associated factor. Base is the factor that is used as the comparison (coef. = 0).
^^ Chi-Square response was intra-site differences grouped into high (>5 ppb) and low (≤5 ppb) bands, with the p-value (χ^2) reported.
^^^ Chi-Square test was run without the factor roof for siting due to only one example within this category.
* p-Value < 0.05.
** p-Value < 0.01.

Table 3
Intra-site summary statistics for the low-cost sites. MAE – Mean Absolute Error; RMSE – Root Mean Square Error; R^2 – correlation coefficient; $\rho_{X,Y}$ – Spearman rank correlation.

Site	MAE (ppb)		RMSE (ppb)		R^2 (%)		$\rho_{X,Y}$ (%)		Cohen's *d*	
	1-h	24-h	1-h	24-h	1-h	24-h	1-h	24-h	1-h	24-h
a	1.73	1.55	2.15	1.85	0.77	0.84	0.97	0.95	0.01	0.01
b	4.1	3.89	5.02	4.74	0.42	0.21	0.65	0.46	0.07	0.03
c	3.3	2.44	4.3	3.12	0.54	0.5	0.73	0.63	0.12	0.11
d	4.65	4.12	5.92	5.17	0.45	0.39	0.69	0.63	0.04	0.04
e	3.34	3.65	4.91	5.3	0.4	0.4	0.79	0.77	0.31	0.42
f	4.2	4	5.44	5.09	0.43	0.3	0.66	0.51	0.32	0.41
g	4.58	3.8	6.39	5.06	0.54	0.69	0.77	0.82	0.11	0.13
h	2.97	2.64	3.99	3.48	0.43	0.42	0.74	0.66	0.1	0.14
i	3.55	2.54	4.69	3.42	0.6	0.66	0.78	0.81	0.09	0.1
j	3.69	3.59	4.24	4.04	0.73	0.73	0.89	0.9	0.15	0.1
k	4.99	4.23	6.19	5.08	0.44	0.52	0.73	0.77	0.52	0.67
l	3.12	2.87	3.9	3.56	0.54	0.57	0.81	0.77	0.13	0.15
m	4.15	3.91	5.46	5.3	0.48	0.51	0.77	0.72	0.51	0.58
n	4.39	4.36	5.25	5.13	0.5	0.39	0.7	0.63	0.32	0.38

The second example examined effects of delivery and collection of children from school using motor vehicles (school runs), as one site was located within a primary school and had co-located wind data. This activity covers a longer period of time than lawn-mowing and often has no obvious spike, however, it may impact standard reporting times from the localized activity within the school grounds, and contain a seasonal component due to holidays. No comparable nearby site was available to analyze the spatial impact. Data was filtered to periods where wind direction was from the car park/pick-up area and during school hours (0700–1600 Monday–Friday), which was then grouped into school term or school holiday periods in order to control for potential confounding effects due to time of day or day of week. Assessment between the two groups was compared on their one and 8 haverages to again check for differences at typical reporting scales.

2.3. Inter-site analysis methods

2.3.1. Inter-site differences

Inter-site variability was carried out in two ways. The first method was by comparing each low-cost site (using the average of the two instruments where both data had been validated) to the analyzer run year-round at Patumahoe using similar statistical tests to the intra-site analysis for network performance. Variation in correlation to the analyzer of the different sites was used to uncover new information about O_3 patterns across the city. This analysis was supplemented by using hierarchical cluster analysis among all of the sites (both analyzer and low-cost) to see if any similarities or clusters were present. Sites were clustered by their similarities using hourly median O_3 Spearman rank correlations against each of the other sites (with the ranked correlation used to minimize the impact of outliers). Clusters were formed using the complete linkage method, which is derived from selecting the smallest maximum distance, or dissimilarity, among groups, with convergence reached once all groups were within one cluster. Sites where larger dissimilarities were found (greater O_3 variability and poor correlation) illustrated how a denser network can add new information (along with potential redundancies where dissimilarity was low). This combination

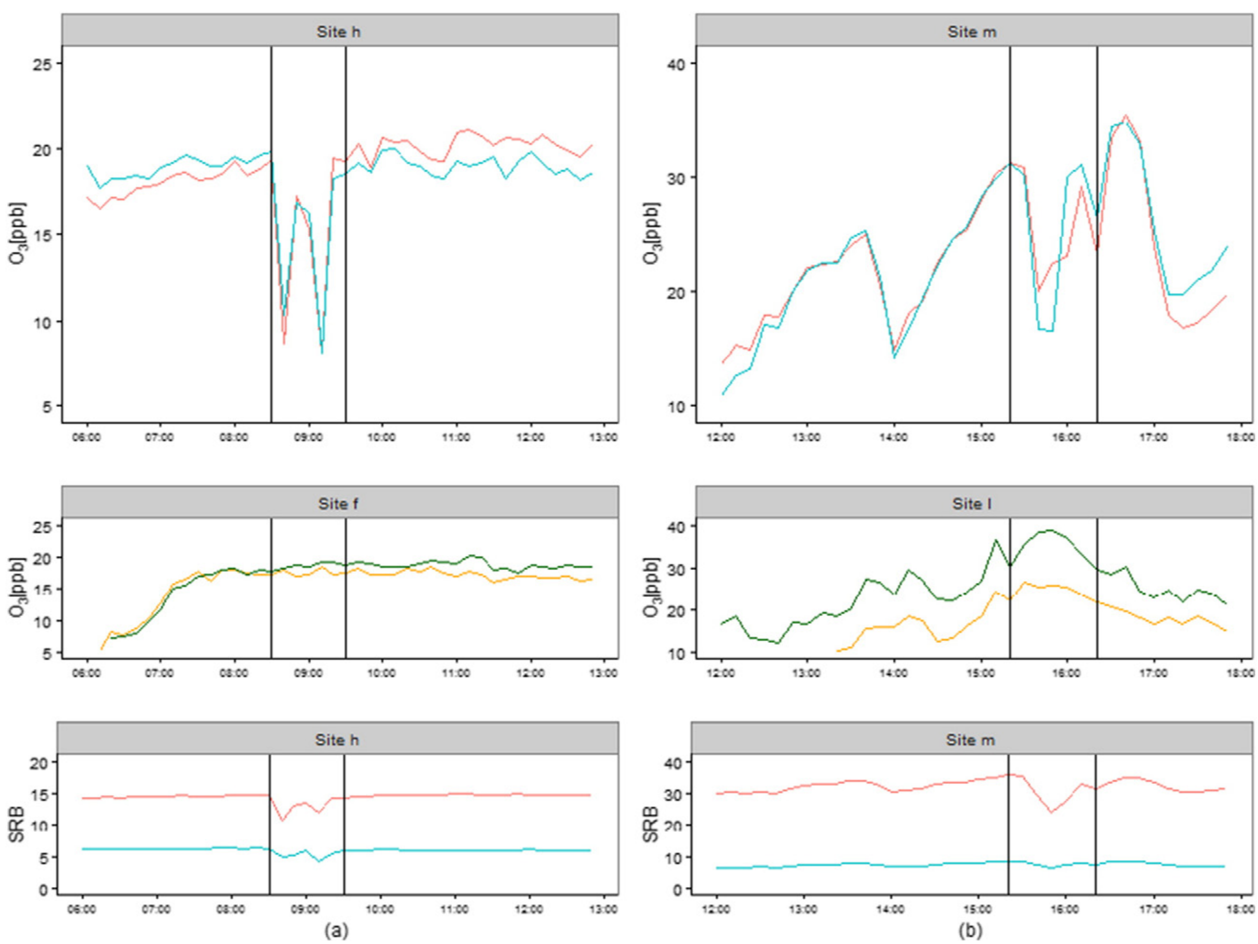

Fig. 4. 10-min ozone and sensor baseline resistance (SRB) data for two sites, h and m, during the day of mowing, with comparison to O_3 at sites f and l in proximity. (a) is for the Site h lawn-mowing event and (b) is for Site m. The outlined period is where lawn-mowing was actively noted. The colored lines represent the two individual instrument data at each site.

of methods was chosen to enrich understanding of any long-term spatial patterns across the city.

2.3.2. Regional reporting

Evaluation of the usefulness of data from uncontrolled site types at representing regional concentrations was completed by comparison of a select number of sites in close proximity to a nearby reference analyzer (<20 km). Two reference analyzers were compared (Musick Point and Whangaparaoa), as no low-cost sites were within close proximity to Patumahoe (Fig. 1). Locations and installations are shown in Fig. 2, with all breaching standard requirements for siting when using analyzer station specifications. Some differences in concentrations among sites can be expected due to local-scale effects and due to natural variation in the atmosphere, but should be relatively minor within the network objective and for such distances (e.g. indicative exposure impacts on the regularly patterned pollutant O_3). Data were compared using similar performance statistics as the intra- and inter-analysis, along with correlation plots on the hourly and rolling eight-hourly data. The objective was to explore how useful uncontrolled sites can be for reporting on wider pollution patterns.

3. Results

3.1. Intra-site differences

The variability of the intra-site mean absolute difference was <5 ppb for most locations (Fig. 3) which showed good agreement between instruments on O_3 concentration within a site. Length of time did not appear to determine instrument differences, and therefore the changes could not be attributed to sensor ageing or drift alone. This allowed one to make comparisons among instrument concentration differences without adjusting for length of time. Some of the smallest differences (MAE: <2 ppb) were found at Sites *a*, *e*, *h* and *j*, all uncontrolled sites (two in university grounds, two in private homes, Fig. 3; Table 1). It appeared that the site type did not affect the magnitude of the difference between co-located sensors with a two-sample *t*-test finding the uncontrolled sites having smaller mean absolute differences (p-value ~ 0, controlled MAE = 4.03 ppb, uncontrolled MAE = 3.91 ppb).

Absolute differences between co-located sensors were compared against explanatory variables using linear regression to see if failure could be associated with particular siting variables. Data required log-transformation due to normality assumptions when using regression. Ground cover (e.g. grass), awnings, land-use, and distance to a water body were not found to be associated with intra-site differences, neither was proximity to a small emissions source, such as a single lane driveway. Larger intra-site variability was noted where instruments were east-facing (and therefore often having longer direct sun exposure time) and where there were nearby large emission sources such as a car-park (Table 2). We suspect that high temperatures within the plastic instrument enclosures due to direct sun exposure would affect the accuracy of the measurement circuits in the instrument and may also result in the release of hydrocarbons from internal plastic components. These could then be drawn into the sensor housing where it would then react with ozone to create instrument-specific errors. Along with this, if inlet tubes or filters became hot then O_3 may have decomposed before reaching the sensor. Intra-site variability also appeared sensitive to some types of sensor placement. Higher variability was noted when the sensors were on roofs due to higher exposure to the elements, especially direct sun (however this option may be appealing for aesthetic reasons as out of view). Instruments that were back to back also recorded larger differences, possibly also due to uneven sun exposure. Types of instrument mounting or locations (either on walls, on reference stations, or on balconies) was not found to be significantly different (χ^2 p-value = 0.21), giving evidence that the different mounting options did not impact the size of concentration differences. Site *a*, located on a wooden shed wall at an agricultural site in proximity to a small

Table 4
Inter-site summary statistics for the low-cost sites against the Patumahoe reference station.

	MAE (ppb)		RMSE (ppb)		R^2 (%)		$\rho_{X,Y}$ (%)		Cohen's *d*	
	1-h	8-h	1-h	8-h	1-h	8-h	1-h	8-h	1-h	8-h
a	5.07	4.45	6.7	5.76	0.45	0.55	0.72	0.78	0.52	0.58
b	5.42	4.71	6.85	5.92	0.32	0.38	0.6	0.63	0.29	0.32
c	6.57	4.7	6.1	5.33	0.52	0.61	0.75	0.79	1.12	1.21
d	8.16	7.89	9.74	9.26	0.4	0.47	0.69	0.72	0.9	0.97
e	5.38	4.56	6.98	5.9	0.16	0.25	0.49	0.56	0.1	0.11
f	8.42	8.04	10.18	9.61	0.22	0.26	0.5	0.52	0.93	1.01
g	7.04	6.41	8.96	8.03	0.23	0.33	0.57	0.63	0.46	0.49
h	6.01	5.46	7.58	6.89	0.28	0.33	0.57	0.59	0.46	0.92
i	3.69	3.07	4.83	4.05	0.57	0.65	0.77	0.81	0.12	0.15
j	5.28	4.57	6.88	5.88	0.39	0.47	0.66	0.7	0.24	0.27
k	4.63	3.85	5.63	4.67	0.37	0.49	0.68	0.74	0.17	0.19
l	6.85	6.27	8.38	7.61	0.2	0.27	0.52	0.56	0.69	0.76
m	6.57	5.8	7.96	7.04	0.02	0.02	0.16	0.16	0.89	0.97
n	4.18	3.31	5.31	4.19	0.42	0.56	0.72	0.79	0.05	0.06

airfield, had the best results for all summary statistics, with good precision and accuracy and low practical difference between sensors (Table 3). This was surprising due to the regular lawn-mowing and agricultural activities, the isolation of the site and so tendency of spiders to make webs that could cover the inlets, and building material type; however, the site had low sun exposure and the instruments were side-by-side. Results therefore suggest that the best strategy for capturing a site's concentration effectively can be achieved by placing sensors side-by-side, away from large emission sources, and on surfaces which are not exposed to direct sunlight (e.g. roof surfaces) as instruments appeared to be sensitive to high temperatures.

3.2. Local scale impacts

Data for the two separate lawn-mowing periods are presented in Fig. 4 at ten-minute averages for the two sites, along with their respective sensor baseline resistance and O_3 concentration at a site in close proximity (<10 km).

The data for Site *h* was on a Thursday where the lawn was mowed during 0830–0930 and the data for Site *m* was on a Sunday during 1520–1420 (NZDT). Distinct local-scale patterns can be observed at each site, along with differences between the sites in proximity, which can be due to local-scale activities that have not been noted (e.g. Site *h*, in proximity to a drive-way serving a number of houses, often showed a dip in O_3 concentration around 1500). The spike in O_3 concentrations was observed in both instruments at both sites, and the instruments returned to similar concentrations following this event. This was also captured in the sensor resistance baseline for the four instruments with sudden negative spikes. These attributes were not observed in the nearest sites during the same time, and so results show that extremely local-scale high-impact activities such as lawn-mowing can be picked up by instruments, and that the sensors can return to similar concentrations following such exposure. The impact on a site's one-hour concentration (where sensors were averaged) were minor, with differences of <3.3 ppb for both sites compared to when the lawn-mowing impacts were removed and then averaged (and were <1 ppb difference on the eight-hour averages). The hourly differences between the site where lawn mowing occurred and the site in proximity were <2.5 ppb when lawn mowing period data were removed (and <1 ppb when lawn-mowing data were intact). Due to the distinct characteristic of such events, invalidation of data or alerts if one wanted to improve data quality further could be set up using baseline diagnostics specific to each sensor, such as $(srb_2 - srb_1)/(t_2 - t_1) > \alpha$ where α is some arbitrarily set threshold to note changes between times *1* and *2*, could be instated so that active observation of activities are not required.

The second example of local-scale impacts analyzed data from Site *b*, located on school grounds, to investigate the impact of the school runs. Results (in Supplementary material) showed consistent differences in hourly averages when data were grouped into either school term or school holiday periods and filtered to school hour times and wind direction from the carpark area. Overall, concentrations were lower and less variable during the school term ($n = 178$ h), with median (inter-quartile range, IQR) of 15.4 (11.2, 17.8) ppb compared to 16.3 (11.4, 22.7) ppb during the school holidays ($n = 88$ h), with a statistically significant difference between the medians (p-value $= 0.01$). However, this difference would appear to have low practical significance due to the large overlap of the IQR. The start (0700) and finish (1600) of the monitored period returned to similar concentrations for each group (<1 ppb), giving support that the presence of cars from the school-run was causing this difference, along with larger dips during the more common drop-off (0900) and pick-up (1500) times (4 ppb difference for both times). Differences in the eight-hour averages were negligible (1 ppb) due to the effect of rolling hours, although a slight dip during the day could be observed. This result showed the small, but significant, impact of school runs on the surrounding environment, providing useful information on siting from a network analysis perspective.

3.3. Inter-site differences

The comparison between the averaged low-cost sites and the Patumahoe analyzer concentrations showed reasonable O_3 variability across Auckland (Table 4).

Sites with low MAE and RMSE scores (similar concentrations) were Sites *i*, *k* and *n*, which were all located in semi-rural areas (similar to the analyzer setting). Eight and nine of the 14 sites had hourly and eight-hourly ranked correlations ≥60% to Patumahoe respectively, showing that overall, concentration rankings were similar between the analyzer and the sites (good agreement on high and low concentration periods). The correlation coefficient (R^2) between the sites and the analyzer were typically poor however, which illustrates the degree of O_3 variability across the city. R^2 values are highly impacted by outliers, which here can represent local-scale events, and so the results supported the value in measuring at more sites. Sites with relatively high R^2 (and so where the analyzer was able to explain a higher degree of the observed variability) were sites *a*, *c*, *i*, and *n*, which were not downwind of the city center (although in different directions to one another and up to 92 km away). The city center (an area creating large amounts of precursor emissions) and prevailing wind direction therefore seem important in explaining regional O_3 variability, as distance alone was not found to be associated with site variability (Table 4). Hierarchical cluster analysis further supported this conclusion (Fig. 5), where sites were grouped by

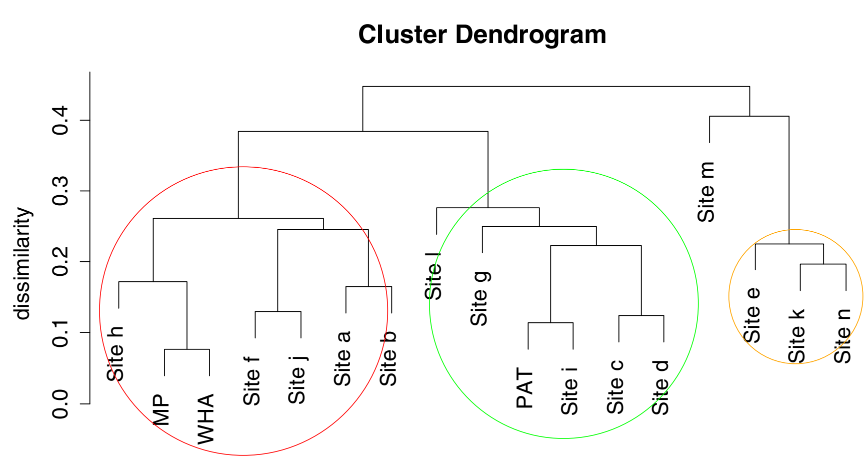

Fig. 5. Dendrogram for the low-cost sites using the complete linkage method and Spearman correlation coefficient as the indicator. Three groups were identified (red = central; green = west; orange = east).

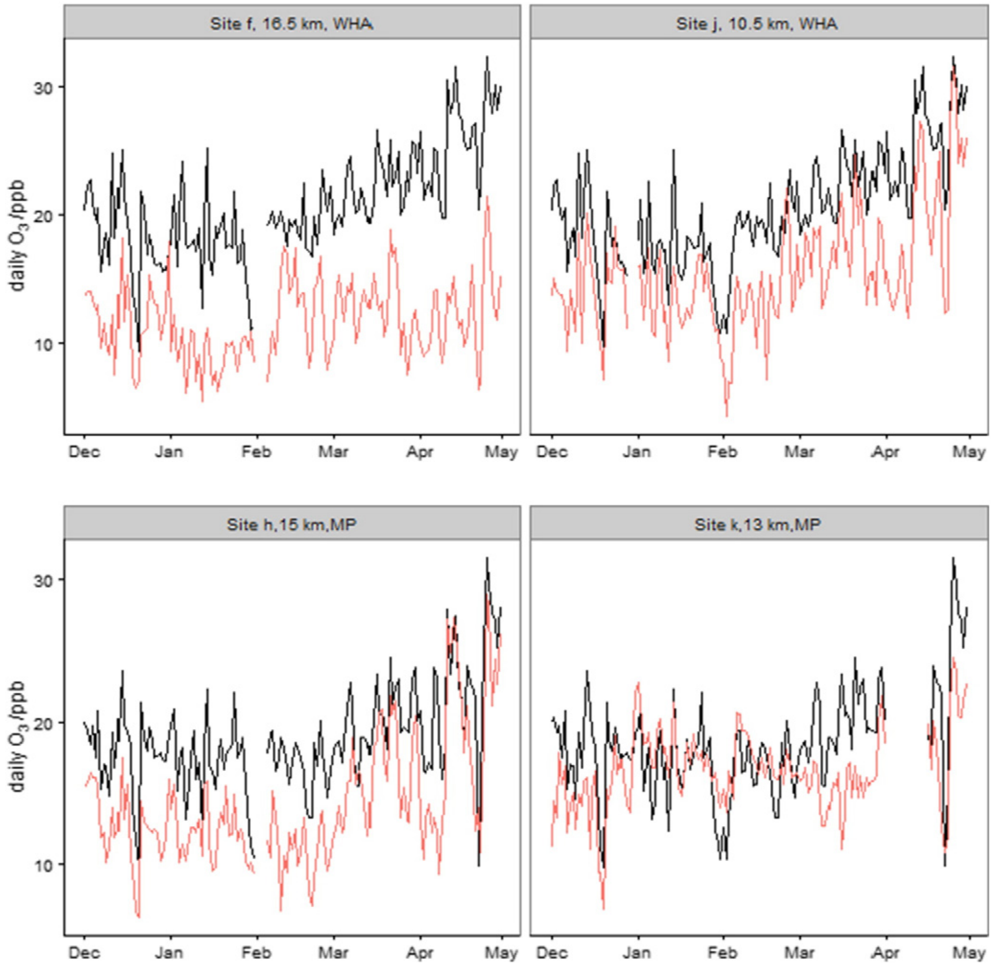

Fig. 6. Daily time-series for the four low-cost sensors (red) in close proximity (<20 km) to reference analyzers (black). Distance (km) and specific analyzer are identified in sub-plot titles.

their ranked correlation similarities among each of the sites (for both low-cost and analyzer). Those sites that were linked at a lower dissimilarity (which here is 1-ρ) were those that had small physical distances from one another (e.g. Sites *c* and *d*), similar direction towards the urban center (e.g. Sites *a* and *b*), and comparable land-use characteristics (e.g. Sites *f* and *j*) (Table 1; Fig. 5). Three predominant clusters were found, the first being for Sites *c*, *d*, *g*, *i* (*PAT*), and *l*, which were towards the west of the city center, the second being for Sites *e*, *k* and *n*, which were towards the east, and the third being for Sites *a*, *b*, *f*, *h*, *j*, *MP*, and *WHA*, which were nearby and around the city center. Site *m* was found to be the most different to all the groups, although was more similar to the second, which was unusual based on the close physical proximity to the other cluster sites (Fig. 1). The setting was at a high elevation (over 300 m), and so may be a reason for this discrepancy. This may show the influence of synoptic meteorological parameters and topography/height upon O_3 distribution and patterns, although not enough sites at high heights were available to test these parameters further. In relation to sensor mounting effects, no obvious cluster was observed among the different siting types. This helps to support the finding that specific instrument mounting had minimal impact on the resulting data.

3.4. Regional reporting

The data from a subset of instruments in crowd-sourced locations (Fig. 2) were analyzed against analyzer data within close proximity to check for consistency in O_3 reporting and their subsequent use in explaining concentrations for a wider area. The instruments were found to have similar time-series patterns and most had agreeable correlation plots to the analyzers, albeit with different magnitudes and variability (Fig. 6). Both examined analyzers are located on the ends of peninsulas and have high elevation or inlets (83 m elevation and 12 m inlet, Auckland Regional Council, 2005), which could explain the often suppressed diurnal cycle observed (Adeeb and Shooter, 2004). Site *f*, located in a residential setting, had the biggest difference to the compared analyzer (Table 5). This may be due to the limited O_3 range measured at the low-cost site, which may be due to the presence of titration emissions (e.g. traffic-related NO_X). Site *k* concentrations were similar to the analyzer, although high scatter was observed, causing low R^2 results (potential local-scale short-term effects). Effect sizes, however, were considered to be acceptable for three of the sites; that is, the observed patterns had 'medium' or below ($d < 0.8$) practical differences when using the Cohen's *d* statistic and widely-adopted thresholds (Cohen, 1988). This meant that three of the crowd-sourced sites were providing

Table 5
Inter-site summary statistics for the four sites against the reference analyzer in proximity (<20 km) at hourly, eight, and daily resolution.

Site	MAE (ppb)		RMSE (ppb)		R^2 (%)		$\rho_{X,Y}$ (%)		Cohen's *d*	
	1-h	8-h	1-h	8-h	1-h	8-h	1-h	8-h	1-h	8-h
f	9.2	9.25	10.57	10.41	0.23	0.24	0.48	0.49	1.32	1.38
h	5.28	4.98	6.42	5.87	0.47	0.55	0.67	0.73	0.74	0.82
j	5.29	5.12	6.86	6.27	0.57	0.62	0.74	0.76	0.71	0.78
k	4.02	3.67	5.16	4.64	0.28	0.33	0.51	0.53	0.38	0.43

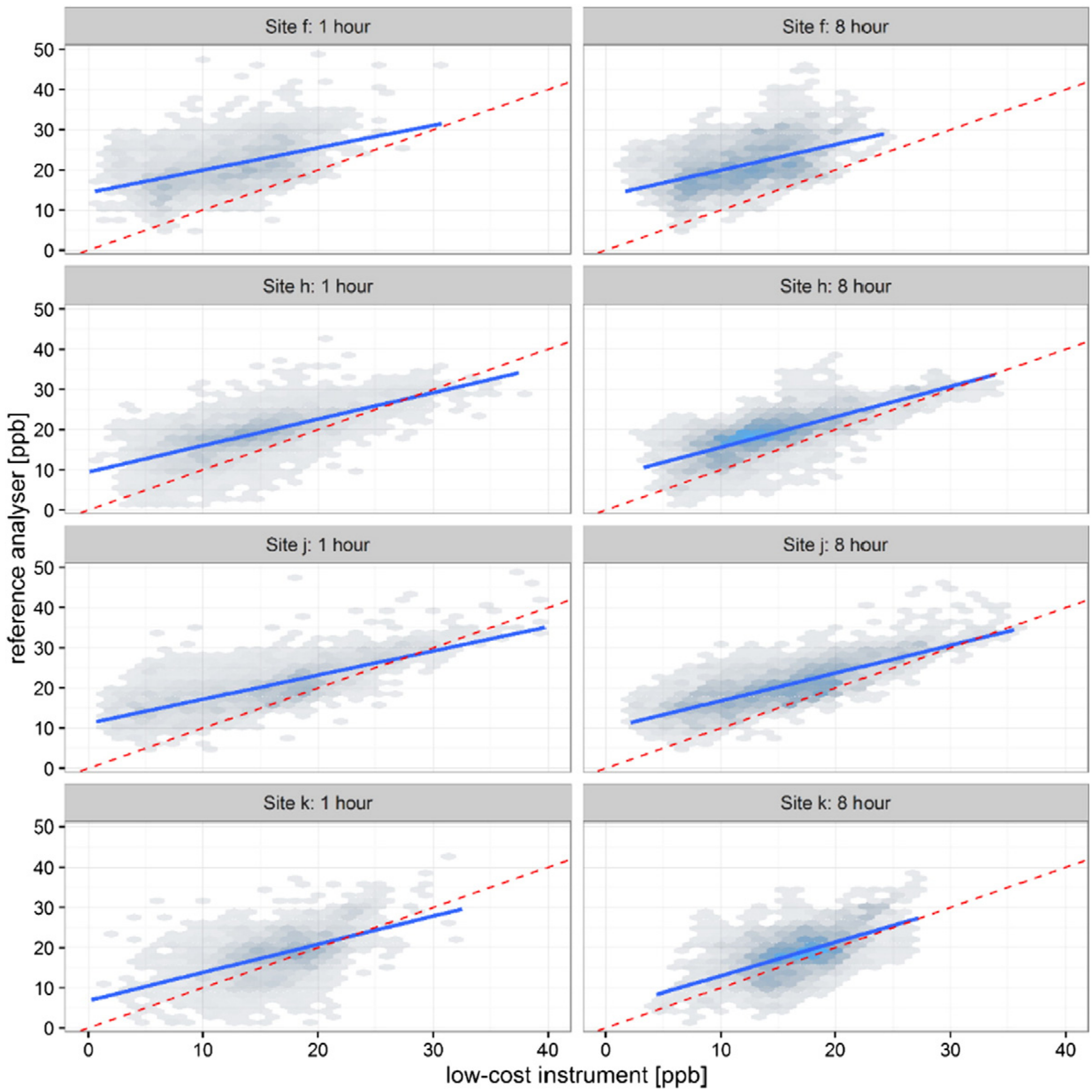

Fig. 7. hex-bin plots for the subset of reference and low-cost sites in proximity at one-hour (left) and eight-hour (right) resolution. The red dashed line denotes the 1:1 fit and the blue solid line denotes the least-squares fit.

practically similar concentrations to the analyzer in proximity (and therefore could be used as indicators of O_3 for a wider area).

Comparisons between the analyzer and low-cost instruments were made by hex-plots, which showed the crowd-sourced sites to often record lower concentrations (Fig. 7). This result appeared sensible for O_3, as analyzers are often placed in locations where concentrations are assumed high and are free of any chemical interferences. Three of the sites here were on private homes, where concentrations can be assumed to be lower due to local-scale activities that impact and reduce concentrations (e.g. traffic-related). The fourth site, Site *k*, was within a vineyard (arguably similar surroundings and controls to parklands) and downwind of the city center, and so it would be believable to have similar concentrations to the analyzer (in parklands and downwind).

4. Conclusion

Data from low-cost instruments can add interesting and valuable information for an area through improved spatial resolution and through highlighting relationships among sites. Different siting types and local-scale effects did not appear to have significant impacts on monitored O_3, with no clear evidence that siting caused large differences between the two co-located sensors, and that short-term activities did not impact longer term results at which reporting is made. Spatial variability was often low within and high between sites, which provided confidence that the observed differences were real and not a type of instrument functioning concern (other than keeping the instruments shaded). Crowd-sourced datasets appeared capable of capturing wider concentration trends, and therefore be representative for regional concentrations. Land-use descriptions, direction towards the urban center, and distance among sites appeared to be more important in determining patterns than specific siting details, with clusters having similar characteristics (and no obvious cluster based on instrument mounting). The observations made here may help to alleviate concerns about instrument mounting effects for crowd-sourced monitoring networks.

Acknowledgements

We would like to thank the individuals and facilities who generously hosted instruments as well as the research assistants who helped with data sorting and fieldwork. We are also grateful to Auckland Council for the provision of data and access to sites. Funding was provided by a Callaghan Innovation Education Fellowship, Contract Number: AIRQU1301/34810, MacDiarmid Institute for Advanced Materials and Nanotechnology, NZ Ministry of Business, Innovation and Employment, contract UOAX 1413, and The University of Auckland Faculty Research Development Fund (awarded to J. Salmond) (UoA FRDF 3704143). We would also like to thank the two anonymous reviewers for comments on improving this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.scitotenv.2016.09.177.

References

Adeeb, F., Shooter, D., 2004. Variation of surface ozone in the ambient air of Auckland, New Zealand. Environ. Monit. Assess. 95 (1–3), 201–220.

Alavi-Shoshtari, M., Williams, D.E., Salmond, J.A., Kaipio, J.P., 2013. Detection of malfunctions in sensor networks. Environmetrics 24 (4), 227–236.

Auckland Regional Council, 2005. The Ambient Air Quality Monitoring Network in the Auckland Region TP296. Auckland Regional Council, Auckland.

Bart, M., Williams, D.E., Ainslie, B., McKendry, I., Salmond, J., Grange, S.K., et al., 2014. High density ozone monitoring using gas sensitive semi-conductor sensors in the Lower Fraser Valley, British Columbia. Environ. Sci. Technol. 48 (7), 3970–3977.

Bonney, R., Phillips, T.B., Ballard, H.L., Enck, J.W., 2016. Can citizen science enhance public understanding of science? Public Underst. Sci. 25 (1), 2–16.

Brienza, S., Galli, A., Anastasi, G., Bruschi, P., 2015. A low-cost sensing system for cooperative air quality monitoring in urban areas. Sensors 15 (6), 12242–12259.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition. Lawrence Erlbaum, Hillside, N.J.

Deville Cavellin, L., Weichenthal, S., Tack, R., Ragettli, M., Smargiassi, A., Hatzopoulou, M., 2016. Investigating the use of portable air pollution sensors to capture the spatial variability of traffic-related air pollution. Environ. Sci. Technol. 50 (1), 313–320.

Ho, H.C., Knudby, A., Sirovyak, P., Xu, Y., Hodul, M., Henderson, S.B., 2014. Mapping maximum urban air temperature on hot summer days. Remote Sens. Environ. 154, 38–45.

Jiang, N., Dirks, K.N., Luo, K., 2014. Effects of local, synoptic and large-scale climate conditions on daily nitrogen dioxide concentrations in Auckland, New Zealand. Int. J. Climatol. 34 (6), 1883–1897.

Lin, C., Gillespie, J., Schuder, M.D., Duberstein, W., Beverland, I.J., Heal, M.R., 2015. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. Atmos. Environ. 100, 111–116.

Ministry for the Environment, 2009. Good Practice Guide for Air Quality Monitoring and Data Management 2009. Ministry for the Environment, Wellington.

Miskell, G., Salmond, J.A., Alavi-Shoshtari, M., Bart, M., Ainslie, B., Grange, S.K., et al., 2016. Data verification tools for minimizing management costs of dense air-quality monitoring networks. Environ. Sci. Technol. 50 (2), 835–846.

Moosavi, V., Aschwanden, G., Velasco, E., 2015. Finding candidate locations for aerosol pollution monitoring at street level using a data-driven methodology. Atmos. Meas. Tech. 8 (9), 3563–3575.

Nieuwenhuijsen, M.J., Donaire-Gonzalez, D., Rivas, I., De Castro, M., Cirach, M., Hoek, G., et al., 2015. Variability in and agreement between modelled and personal continuously measured black carbon levels using novel smartphone and sensor technologies. Environ. Sci. Technol. 49 (5), 2977–2982.

Ottinger, G., 2010. Buckets of resistance: standards and the effectiveness of citizen science. Sci. Technol. Hum. Values 35 (2), 244–270.

SCAMD, 2015. Air Quality Sensor Performance Evaluation Center. Retrieved from http://www.aqmd.gov/aq-spec/.

Snyder, E., Watkins, T., Solomon, P., Thoma, E., Williams, R.W., Hagler, G.S.W., et al., 2013. The changing paradigm of air pollution monitoring. Environ. Sci. Technol. 47 (20), 11369–11377.

Tregidgo, D.J., West, S.E., Ashmore, M.R., 2013. Can citizen science produce good science? Testing the OPAL Air Survey methodology, using lichens as indicators of nitrogenous pollution. Environ. Pollut. 182, 448–451.

U.S. Environmental Protection Agency, 2013. QA Handbook for Air Pollution Measurement Systems, Volume 2, Ambient A]ir Quality Monitoring Program. Retrieved from https://www3.epa.gov/ttnamti1/files/ambient/pm25/qa/QA-Handbook-Vol-II.pdf.

U.S. Environmental Protection Agency (n.d.). Next Generation Air Measuring Research. (Retrieved from) https://www.epa.gov/air-research/next-generation-air-measuring-research

Williams, D.E., Henshaw, G.S., Bart, M., Laing, G., Wagner, J., Naisbitt, S., et al., 2013. Validation of low-cost ozone measurement instruments suitable for use in an air-quality monitoring network. Meas. Sci. Technol. 24 (6).

Wolters, E.A., Steel, B.S., Lach, D., Kloepfer, D., 2016. What is the best available science? A comparison of marine scientists, managers, and interest groups in the United States. Ocean Coast. Manag. 122, 95–102.

Xiang, Y., Tang, Y., Zhu, W., 2016. Mobile sensor network noise reduction and recalibration using a Bayesian network. Atmos. Meas. Tech. 9 (2), 347–357.