Low-cost sensors and microscale land use regression: data fusion to resolve air quality variations with high spatial and temporal resolution

L.F. Weissert[1*], K. Alberti[2], G. Miskell[1], W. Pattinson[3], J.A. Salmond[4], G. Henshaw.[2], David E. Williams[1,5]

1. School of Chemical Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
2. Aeroqual Ltd, 460 Rosebank Rd, Avondale, Auckland 1026, New Zealand
3. Mote, 40 George St, Mount Eden, Auckland 1024, New Zealand
4. School of Environments, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
5. MacDiarmid Institute for Advanced Materials and Nanotechnology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

*Corresponding author: l.weissert@auckland.ac.nz; david.williams@auckland.ac.nz

Keywords: Land-use regression; Spatio-temporal model; nitrogen dioxide; ozone; Air pollution;

Abstract

The strong temporal and spatial gradients in NO_2 concentrations frequently observed in urban microenvironments are very difficult to measure and model accurately. Recent developments in low-cost air quality instruments have led to improvement in the spatial coverage of time-resolved measurement, however interpolation is still needed to map pollutant concentrations and connect time- as well as space- dependent variations to urban design features. Here we propose a novel approach that uses a previously-described microscale land use regression (LUR) model to spatially interpolate data from a well-calibrated network of low-cost air quality instruments. We use a semiconducting oxide-based ozone sensor to provide a robust correction of the output of an electrochemical NO_2 sensor for ozone interference. We characterise signal noise probably associated with meniscus fluctuations as a significant error source, that can be handled with appropriate signal averaging. The LUR model is used to provide high spatial resolution in the data set, whilst correlation with sensor measurements provides a time-dependent estimate associated with different land use types. Observations from the network of instruments showed marked variability in NO_2 concentrations over short distances (on the scale of 100 m), with highest concentrations reached near bus stops, intersections and under shop awnings. This approach connects the complex time- and space-dependent variations to urban design features and is a promising way forward as a basis for objective spatial mapping of time-dependent mean concentration fields and local population exposure estimates.

1. Introduction

Epidemiological studies often use modelling techniques, such as land use regression modelling or dispersion modelling, to quantify pollutant concentrations (Hoek et al., 2008; Vardoulakis et al., 2011). Land use regression (LUR) models are suitable for mapping air pollution with minimal additional data acquisition and are therefore a low-cost alternative to fixed monitoring (Jerrett et al., 2005). However, standard LUR models lack information about temporal variability particularly at the daily, hourly and sub-hourly scales. Land use regression studies also generally focus on city and regional scales where the explanatory variables have large spatial scale. The spatial resolution of the models is km and not representative of microscale environments (Mukerjee et al., 2012; Weissert et al., 2018). Thus, recent studies have focussed on urban design as a driver of air pollution. 'Microscale' LUR models have been developed that emphasise local urban design variables (Miskell et al., 2015; Weissert et al., 2018). The distance scale of these models is 50 m and, indeed, over such small distance scales the variations are significant. To improve the temporal resolution of LUR models, previous studies have shown some success using temporally dynamic predictor variables (i.e. traffic, meteorology) (Masiol et al., 2018; Miskell et al., 2018b; Son et al., 2018; Yeganeh et al., 2018). This approach may be limited by the availability of such data at a high temporal and spatial resolution. Other studies have linked regulatory monitoring data with the output from LUR models (Cordioli et al., 2017; Johnson et al., 2013; Nethery et al., 2008; Slama et al., 2007). However, regulatory monitoring networks are typically sparsely distributed across urban areas and data from just one or two sites are used to calibrate LUR models (Cordioli et al., 2017; Johnson et al., 2013; Slama et al., 2007). Consequently, these approaches are based on the assumption that temporal variability is the same across large areas (Cordioli et al., 2017; Johnson et al., 2013).

The central question of our research is how to assess the full spatio-temporal pollution field on hourly timescale or shorter, updated at a minimum hourly, and on distance scales of 50 m. Using nitrogen dioxide (NO_2) measurements from a pilot study undertaken in Auckland, New Zealand, we show that results from low-cost instruments have sufficient accuracy and precision to resolve spatial variations, then show how to combine results from a network of low-cost instruments with a microscale LUR model in order to estimate the hourly mean concentration at a high temporal and spatial resolution, taking into account the effect of land use variables. The results provide important insights into site-specific effects on the temporal variability of NO_2 concentrations.

2. Theory

We used a microscale LUR model (Weissert et al., 2018) and a network of low-cost instruments to derive a LUR model that is updated hourly and to identify site-specific effects (e.g. interaction of bus stops and traffic lights at different times of the day).

Let $\bar{c}_{LUR,i}$ denote the long-term time-averaged concentration at site *i* modelled by LUR, using data $[z]$ measured by diffusion tubes. The LUR model is fitted by linear least-squares regression:

$$\mathbb{P}\left(\bar{c}_{LUR,i}\middle|[z]\right) = \mathbb{N}\left(\left[\beta_0 + \sum_j \beta_j q_{j,i}\right], \sigma_{LUR}^2\right) \qquad (1)$$

where $q_{j,i}$ denotes the site characteristics and $[\beta_j]$ are the fitted parameters, which are normally-distributed, independent random variables. Here, $\mathbb{P}$ denotes the probability density and $\mathbb{N}$ denotes the normal distribution ($\mathbb{N}(\mu, \sigma)$). The coefficient β_0 is derived from the β_j and the mean of the data and so is not an independent variable. Equation 1 is a typical LUR formulation as implied in many studies (e.g.Hoek et al., 2008).

We developed a new model for explaining the hourly mean concentration, whose expected value for any particular hour, *l*, and any particular site, *i*, $\bar{c}_i$, is conditional upon the original LUR-modelled concentration for that site, $\bar{c}_{LUR,i}$, the measured set of hourly mean values from the sensors for that hour, [y_k], and the original LUR model applied at the instrument sites, [$\bar{c}_{LUR,k}$]. We assume a linear relationship between the measured, hourly-averaged sensor data, [$y_{k,l}$] for any given hour of any given day, *l*, and the original LUR model:

$$y_{k,l} = \hat{a}_{1,l}\bar{c}_{LUR,k} + e_l \qquad (2)$$

where, provided the sensors and the diffusion tubes have the same calibration slope with respect to the true NO_2 concentrations, any offset error between sensors and diffusion tubes is zero-mean with normally distributed errors. If the model is fitted by linear least squares regression, then e_l are normally-distributed zero-mean, random errors and the slope, $\hat{a}_{1,l}$, is a normally distributed random variable. Then we assume that this model applies at all other sites, *i*. That is, the model for the hourly-averaged concentration at site *i* for any given hour of any given day, *l*, is:

$$\bar{c}_{i,l} = \hat{a}_{1,l}\bar{c}_{LUR,i} + \varepsilon_{i,l} \qquad (3)$$

where the estimates $\hat{a}_{1,l}$ are derived by least-squares fitting of equation 2. The actual distribution of the error term (indication of the model performance), $\varepsilon_{i,l}$ would reflect site-specific effects that are not captured by the original LUR model, and can be determined at the instrument sites, *k*.

3. Materials and Methods

3.1 Location

Auckland is New Zealand's biggest and fastest growing city and is characterised by high traffic volumes with a large proportion of diesel vehicles (Statistics New Zealand, 2013). Road traffic is therefore the largest source of air pollution in Auckland, accounting for almost

120 80% of the NO_x (NO_2 and NO) emissions (Xie et al., 2014). However, Auckland's coastal
121 geographic location facilitates relatively high wind speeds, preventing a large build-up of air
122 pollutants (Senaratne and Shooter, 2004). The focus of this pilot study was on a ~2 km long
123 road section around 4 km south of Auckland's city centre (Figure 1). This section of road
124 (Dominion Road) was chosen as it not only has high vehicle traffic (average daily weekday
125 traffic: 25,901 vehicles) (Auckland Transport, 2017) but is also a busy area for pedestrians
126 using the shops, bars and cafés along the road. Dominion Road is also well-used by buses and
127 commuters to Auckland's airport and its central business district (CBD). We deployed eight
128 instruments at different sides of the road, spaced between ~ 100 m and 1 km apart, and at
129 different distances from intersections (Figure 1a). In addition, in a previous study carried out
130 from November 2016 – February 2017, we measured NO_2 concentrations at 40 sites,
131 representative of the typical land use characteristics along the road section, using Palmes
132 diffusion tubes (Weissert et al., 2018) (Figure 1b). These measurements were used to develop
133 a microscale LUR model for the study area, which is described in detail in Weissert et al.,
134 (2018). The presence of awnings and the number of bus stops within a 100 m buffer had the
135 highest proportional contribution (79% and 17%, respectively) to the modelled NO_2
136 concentrations, with street width and distance to major road playing a minor role. The final
137 model explained 66% of the variability in NO_2 concentrations (RMSE: 3.3 $\mu g\ m^{-3}$) (Weissert
138 et al., 2018). We also had a reference site in Auckland's CBD, 3 km north of Dominion Road
139 (Figure 1c). The study was undertaken over 30 days during summer when mean temperatures
140 ranged between 14°C and 24°C (NIWA, 2017), and the wind was from the SW with speed
141 predominantly in the range 0 – 2 $m\ s^{-1}$ (detailed data in the Supporting Information, SI).

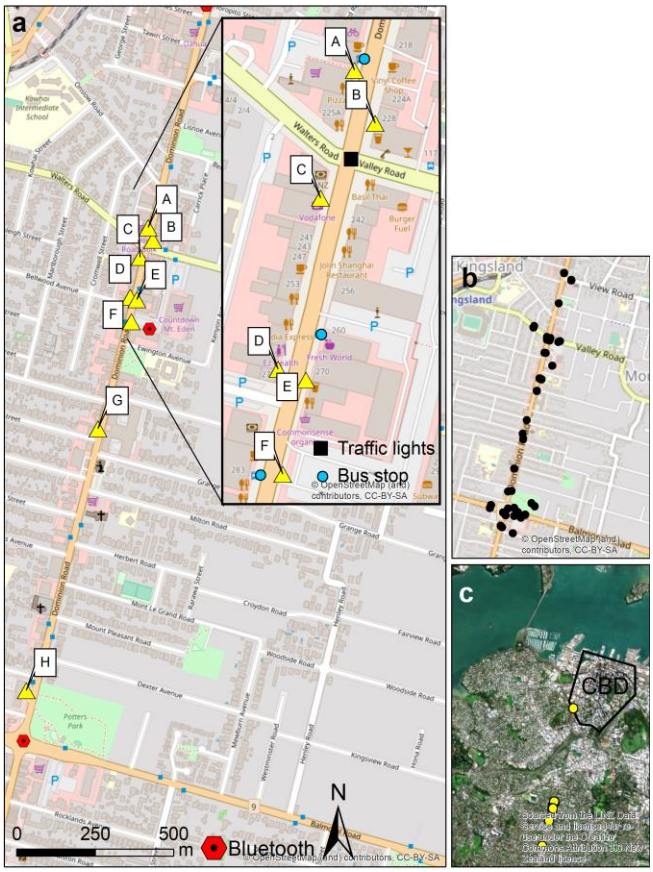


Figure 1. a) Map of instrument locations, b) location of the diffusion tubes deployed in the previous study and c) map of the instrument locations in relation to Auckland's Central Business District (CBD) and the location of the reference site at the edge of the CBD.

3.2 Low-cost instruments

The term 'low-cost sensor' is often used to refer to the assembly of the detection element, measurement electronics, air-inlet, air-sampling and communications systems, and housing and mounting that together deliver the measurement result; 'low-cost' refers to such sensors whose installed capital cost is less than about 2% of that of a regulatory-standard reference instrument. The term 'sensor' can also be taken to mean just the detection element and its specifically associated electronics and calibration. Here, we use it in that way since the whole measurement package (the 'instrument') contains several such sensors. The deployed 'low-cost' instruments were the AQY from Aeroqual Ltd, Auckland, New Zealand. An independent

156 validation of these instruments, for both ozone and NO_2 has also been reported (South Coast
157 AQMD, 2018). Since the accuracy and precision are critical to the assessment of the model
158 developed in this paper, we give a detailed assessment here. Nitrogen dioxide in the 'low-
159 cost' devices was measured using an electrochemical sensor (Membrapore type O3/M5;
160 output ~ 2 nA / (μg m^{-3}) that was exposed directly to the atmosphere hence sampling by
161 diffusion. The 'low-cost' instruments also incorporated an ozone sensor that uses a gas-
162 sensitive semiconducting (GSS) oxide, WO_3, as the detection element (Aliwell et al., 2001;
163 Hansford et al., 2005; Utembe et al., 2006; Williams et al., 2002). Air was drawn through the
164 ozone sensor by a fan, through PTFE tubing and stainless steel filters. The ozone sensor has
165 been extensively validated in both laboratory and field studies (Air Quality Sensor
166 Performance Evaluation Center, 2018; Bart et al., 2014; Deville Cavellin et al., 2016; Lin et
167 al., 2017; Miskell et al., 2018a; Williams et al., 2013). The major cause of drift in these, which
168 occurs over the long term, is particulate deposition in the inlet and on the sensor causing ozone
169 decomposition or diminution in the air flow. The two sensors were mounted side-by-side,
170 protruding from the base of the polycarbonate case and protected by a shield from accidental
171 impacts (photographs in the SI). Sensor signals were sampled every 1-minute, digitised using
172 a 16-bit A-D converter, and transmitted using a cell-phone modem to a remote database. Key
173 features of the device include solar shields to regulate heat, and a sophisticated inlet
174 configuration for the ozone sensor (inert dust filters; anti-static and inert materials).
175
176 **3.2.1 NO_2 measurements: design principles**
177 The NO_2 concentrations averaged over a 2-week period, measured using diffusion tubes in
178 the previous study (Weissert et al., 2018), varied from 4 to 30 μg m^{-3}. Reliable measurement
179 over such a low concentration range, that is sufficiently accurate to discriminate effects
180 between sites, is very challenging using low-cost sensors (Baron and Saffell, 2017; Cordero

et al., 2018; Cross et al., 2017; Feinberg et al., 2018; Lewis et al., 2016; Munir et al., 2019; Popoola et al., 2018) and requires careful consideration of the sensor performance. The electrochemical NO_2 sensor is a conceptually simple but physically complex device (Baron and Saffell, 2017). A current is measured in an electrochemical cell. Conceptually, the electrode potential at a sensing electrode is set such that the reaction rate at the sensing electrode is constrained to a maximum value by the rate of diffusion of the gas through a membrane that covers the electrode. Thus, the current for the reaction of gases at low concentration in air is proportional to the product of gas concentration and diffusion coefficient through the membrane. However, the signal is the sum of an offset current and currents due to the reaction of different gases at the electrodes, and is dependent on electrocatalytic activity of the electrode materials towards the different gases that might be present. The electrode is a porous structure through which gases diffuse to the electrode-electrolyte interface. The reaction current depends on wetted area, gas diffusion through the porous electrode structure, and temperature. Oscillations in water vapour pressure, temperature and water vapour transport can cause oscillations in the meniscus where the liquid electrolyte, electrode and gas meet, hence alterations in wetted area. Sudden changes in humidity indeed cause significant, rapid current transients in these electrochemical gas sensors (Pang et al., 2018; Pang et al., 2017) as we have also observed. The offset current, and the electrode reaction kinetics, can also change with time, since the reactive electrode catalyst materials change slowly with time. These effects can be significant for measurement at typical urban atmosphere concentrations, where the reaction current is typically 2% or less of the offset current for the commercial devices that we used.

Ozone and nitrogen dioxide are both electroreduced at a diffusion-limited rate at the operating potential of the sensor. Thus, the measurement model for the sensor can be written:

206 $$i_{meas} = i_0 - k_1 C_{NO_2} - k_2 C_{O_3} \quad (4)$$

207 $$C_{NO_2} = b_0 - b_1 i_{meas} - b_2 C_{O_3} \quad (5)$$

208

209 Where i_{meas} and i_0 are the measured and the offset current, respectively and the parameters b_0
210 $= i_0/k_1$, $b_1 = 1/k_1$ and $b_2 = k_2/k_1$ have to be determined by calibration. A simple correction can
211 be made if an independent measure of the ozone concentration is available. In the literature,
212 this has been done using a remote reference analyser (Mead et al., 2013) or a model for ozone
213 concentration (Popoola et al., 2018), or another electrochemical sensor (e.g. Cordero et al.,
214 2018; Munir et al., 2019). On the microscale, with traffic-dominated emissions, variability is
215 expected to be large both in time and space, so we rejected a remote or model-based approach.
216 Correction using an electrochemical sensor for ozone relies on constructing sensors for which
217 k_1 and k_2 are sufficiently different; however, since NO_2 and O_3 have virtually the same
218 diffusion coefficient in air, they have almost the same molecular weight, $k_1 \approx k_2$. A
219 disadvantage of that method is that a total of 6 different parameters must be determined, and
220 these have to remain stable, and be confirmed to be stable, over the life of the sensor. In
221 addition, the electrochemical ozone sensor has the sensitivity to temperature and humidity
222 fluctuations that are found for the nitrogen dioxide sensor, making the compensation difficult
223 (Cordero et al., 2018). An alternative is to incorporate an ozone decomposition catalyst on the
224 NO_2 sensor (Hossain et al., 2016). A disadvantage of this is reliance on the unknown
225 efficiency of the ozone catalyst, which must remain high, and be confirmed to remain high,
226 over the life of the sensor. We have adopted an alternative approach, which is to use a
227 semiconducting oxide-based sensor for ozone. As noted above, this sensor has been
228 extensively validated in long-term measurements in the atmosphere, and is negligibly
229 sensitive to humidity changes, to nitrogen oxides at atmosphere concentrations, and to
230 temperature variations (Bart et al., 2014).

3.2.2 NO_2 measurements: errors and validation

Drift or errors in the NO_2 measurement could arise because:

a) The actual value of the offset i_0 could alter from the value stored in the instrument and used in the calculation: background currents in the electrochemical cell are temperature and wetted area (hence relative humidity and interfacial tension) – dependent. Uncalibrated interferences could also cause additional currents and hence an alteration in i_0. A specific issue is the response to rapid humidity changes noted above.

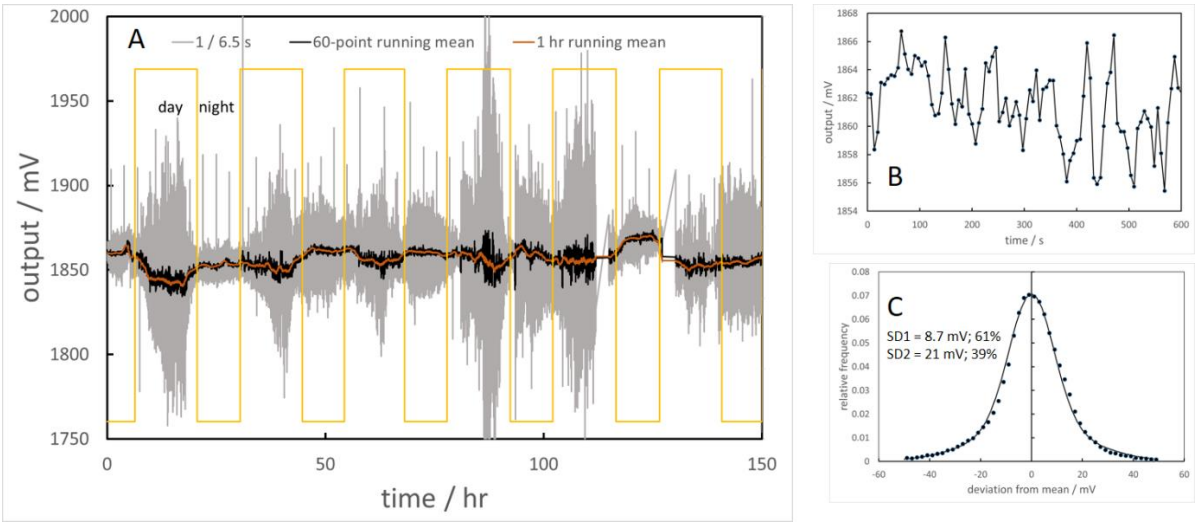
b) The actual value of k_1 could alter from the value stored in the instrument and used in the calculation: clogging of pores in the membrane by dirt deposited in the sensor, or by poisoning or ageing of the electrode.

c) The actual value of b_1 can change from the value stored in the instrument and used in the calculation: decomposition of ozone inside the sensor housing, inlet filter or on the membrane as a consequence of dirt deposition.

d) The instrument-measured value of C_{O3} can drift from the true value: drift in the ozone sensor.

Of particular significance is signal noise that is specifically related to exposure in the atmosphere, previously noted and attributed to atmosphere composition fluctuations (Baron and Saffell, 2017; Mead et al., 2013) (Figure 2a). The fluctuations were of relatively high frequency (Figure 2b). The amplitude distribution of the sampled signal overall could be represented as the sum of two normal distributions, varied with time, and could roughly be classified into ‘noisy’ and ‘quiet’ periods, not correlated with a diurnal cycle (Figure 2c). The standard deviation of the sampled signal did not correlate with the mean signal. This noise

was not present under laboratory conditions of constant humidity (Figure 2d). The noise was also removed if the sensor in the atmosphere was covered with a cap (Figure 2e), or was mounted at the end of a sufficiently long tube, open to the atmosphere at the end remote to the sensor, and could be altered by altering the size of the hole in the cap covering the sensor. Therefore, these fluctuations were not due to the electronics, or to rapid temperature fluctuations but were connected to exposure in the atmosphere and particularly to the design of the sensor inlet and hence probably to fluctuations around the sensor inlet that cause perturbations in the meniscus at the fluid-electrode-gas interface. Such fluctuations could perhaps be related to fluctuations in transport of water between liquid and gas phases. A reliable measurement of (NO_2 + O_3) could be obtained by averaging over a sufficient time. In the case of the instruments used in the present study, we used the hourly average of instantaneous measurements made once per minute. Subsequent instrument designs have used higher frequency sampling, but these were not available for the present study. For the 60-point running mean, the estimated standard deviation in the mean based on the noise amplitude distribution and the sensitivity determined from calibration using the running mean was between 1.7 and 4.3 μg m^{-3} NO_2 ('quiet' and 'noisy' periods respectively) and from the 60-point running standard deviation determined over several days was 2.5 μg m^{-3} NO_2.



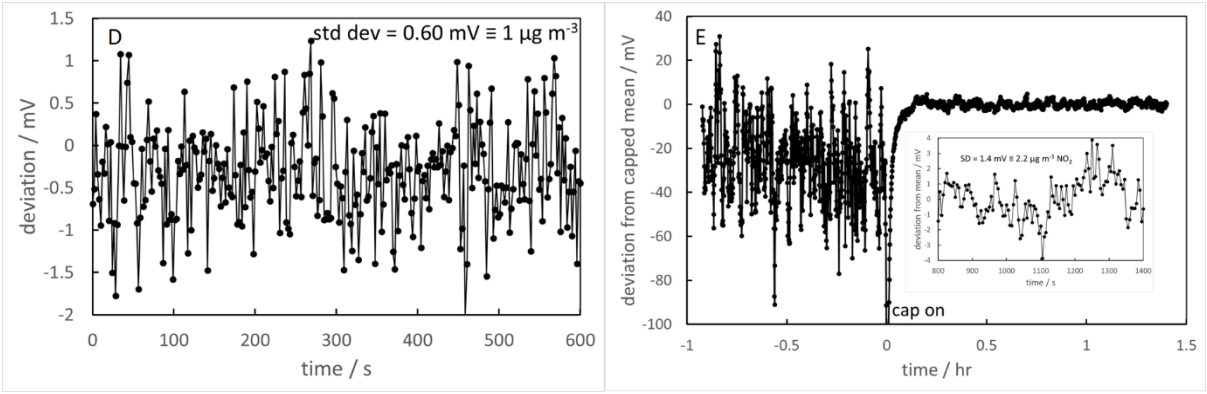


Figure 2. a) Raw electrochemical sensor signal, sampled at 0.15 Hz, illustrating the offset and noise and the effect of signal averaging by a running mean. The cycle of daylight hours is superimposed. b) Expansion of time scale of a segment of the record, to illustrate the high-frequency noise. c) Amplitude distribution of the noise, modelled by a sum of two Gaussians, with standard deviation 8.7 mV (equivalent to 13 µg m^{-3} NO_2) and 21 mV (equivalent to 33 µg m^{-3} NO_2). d) Signal noise at high time resolution in ambient air in a laboratory chamber: signal standard deviation 0.67 mV, equivalent to 1 µg m^{-3} NO_2. e) Effect of placing a cap over the sensor, outside in the atmosphere (inset: segment of the time series with cap on at high time resolution: signal standard deviation 1.4 mV, equivalent to 2.2 µg m^{-3} NO_2.). The instrument output voltage sensitivity to current variations through the sensor is 1.4 mV / nA.

To verify the performance of the devices, first ten devices were co-located for a period of 35 days with reference instruments for O_3 and NO_2 (O_3: Thermo Scientific Model 49i, NO_2: Thermo Scientific Model 42i) at a site close to a motorway junction, approximately 3 km distant from the study area (Figure 1). The performance of the O_3 sensor is critical, since errors propagate directly through to the NO_2 estimation. For hourly-averaged data, individual devices showed excellent correlation between the O_3 sensors and the reference analyser (Adj. R^2 = 0.99) with no drift during the entire co-location period. The O_3 sensors also agreed well among each other, with an Adj. R^2 of 0.98. The major effect on the accuracy of the NO_2

readings was the noise. The results were however satisfactory with sufficient averaging: we chose 1 hr as the time-scale for consistency with practise reporting results from reference stations. As noted above, the standard deviation in the 60-point mean of the raw signal from the NO_2 sensor was equivalent to 2.5 µg m^{-3} NO_2. We determined the measurement model coefficients (equation 5) using the final 10 days of hourly-averaged data from the co-location, before deployment ('calibration period') and evaluated the accuracy and precision relative to the reference instruments using the previous 23 days of data ('test period'). The calibration period was chosen to be immediately prior to deployment to minimise the risk of drift during the deployment period. The test period was of similar length to the site deployment. For the aggregate of all ten NO_2 instruments the root mean square error (RMSE) from the 1:1 line was 5.5 µg m^{-3} (Figure 3). For individual NO_2 instruments, the RMSE evaluated against the reference analyser ranged from 4.1 – 5.1 µg m^{-3} during the test period.

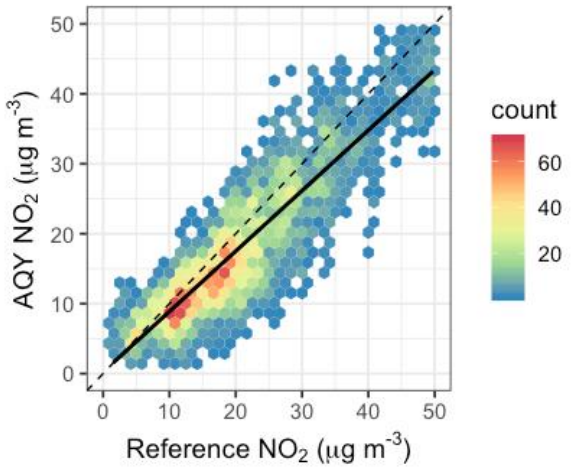


Figure 3. Scatter plot showing the relationship between reference data and the sensor-calculated NO_2 during the test period. Dotted is the 1:1 line; solid is the regression line.

311 To illustrate the stability over time of the results, Figure 4 shows the hourly-average midday
312 NO_2 and O_3 concentration at the reference site for each of the ten sensors, compared with the
313 reference analyser during the test period.
314 After the co-location at the reference site eight of the ten devices were then mounted at
315 different sites across the study area. The instruments were mounted at approximately 3 m
316 height in front of shops and businesses where an external power supply was available. Site
317 photographs are given in the SI. One site (site H) was powered with solar panels and two were
318 battery powered during daytime due to the mains supply being limited by timers. In addition,
319 we measured wind direction and wind speed at site E using a sonic anemometer (WindSonic,
320 Gill Instruments Ltd.) that was installed on a lamp post 5 m above the sidewalk (wind data
321 are shown in the SI).
322 The main concern relating to the data for the deployment period was to demonstrate the
323 stability of the devices over time, and to demonstrate that the observed site-specific effects
324 were significant relative to the errors of measurement. To evaluate the stability of the ozone
325 sensors we used the mean-variance (MV) proxy matching procedure described in previous
326 work (Miskell et al., 2016; Miskell et al., 2018a). There were just two reference instruments
327 available at locations that were not in similar land use to the study site – the motorway junction
328 site and a suburban site approximately 7 km distant, operated by Auckland City Council. We
329 calculated apparent slope and offset using 3-day running mean and variance, as described
330 previously (Miskell et al., 2018a) and did not note any significant drift.
331

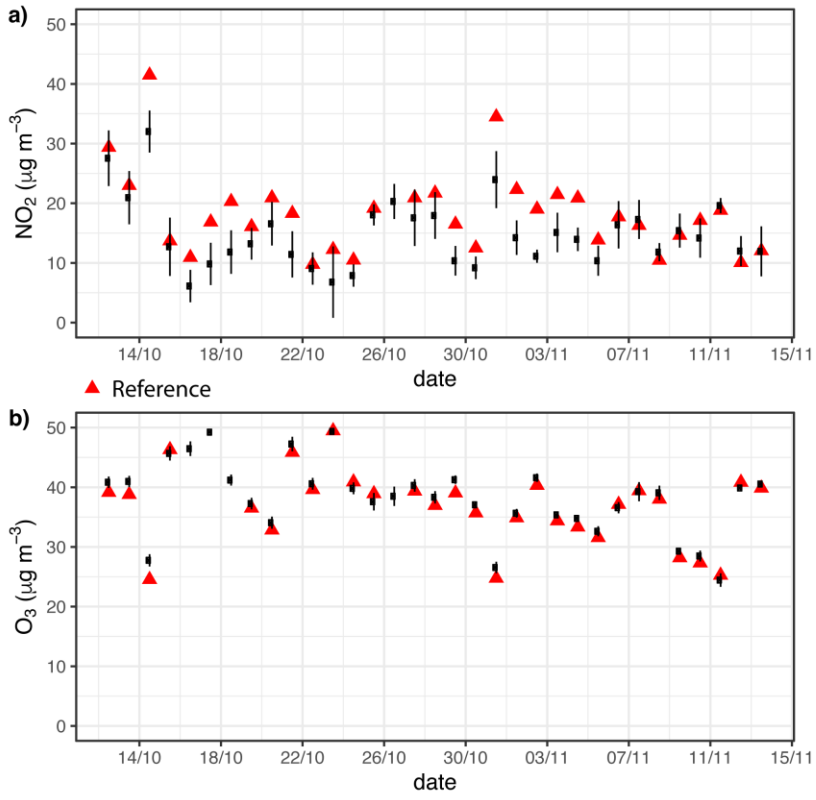
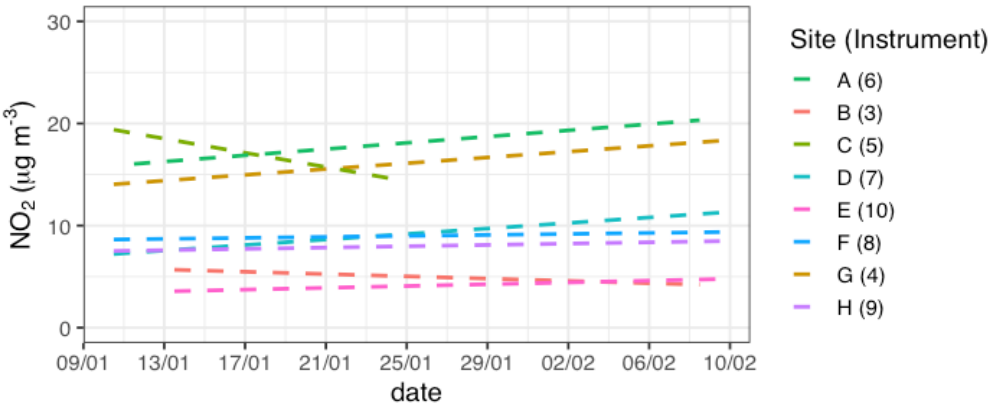


Figure 4. Mean hourly-averaged midday NO_2 concentrations (error bars: standard deviation) for each of the ten instruments at the reference site compared with the reference analyser for the test period.

Figure 4 shows that, at the reference site before the deployment, midday NO_2 did not show any significant long-term trend. Therefore, as a further check to assess the stability of the NO_2 results during the deployment, we simply checked for any long-term trend in the hourly-average values at midday for each site. Figure 5 shows the results for the eight deployment sites. Any trend was at most 5 µg m^{-3} across the whole deployment period.



342

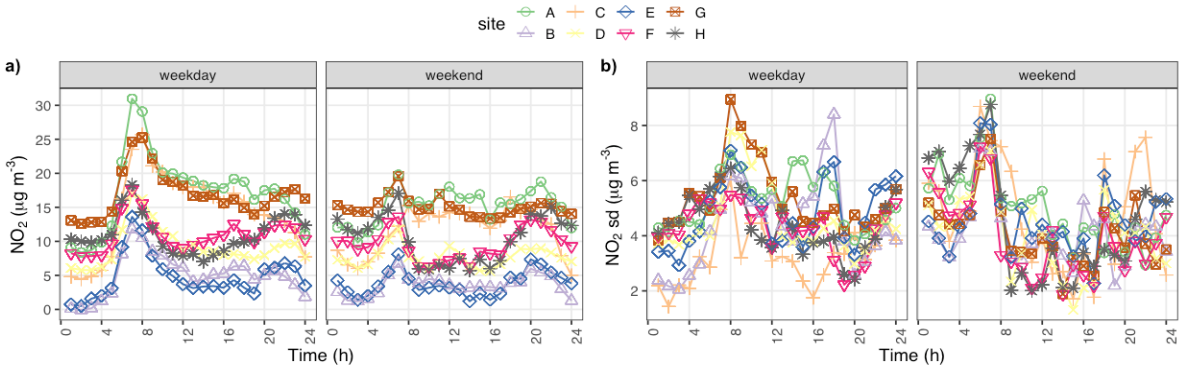
343 Figure 5. Long-term trend for each site (instrument) of measured hourly-averaged midday
344 NO_2 concentration during the deployment period illustrating small drift during the study
345 period.

346

347 **4. Results**

348 **4.1 Spatial and temporal variability of NO_2 concentrations**

349 Results from the low-cost instruments showed very strong gradients in NO_2 concentration in
350 both time, and space, over short distances, of the order of 100 m. Mean diurnal NO_2
351 concentrations are illustrated in Figure 6a, and the variability is further illustrated by the
352 diurnal variation of the standard deviation of measurements over 1 hr (Figure 6b). In general,
353 NO_2 concentrations were low, ranging from around 5 µg m^{-3} at night-time to a maximum
354 between 25-30 µg m^{-3} during the morning rush hour on weekdays, close to heavily-trafficked
355 intersections and bus stops (sites A, C and G: Figure 6). In contrast, other sites, even relatively
356 close by, showed maximum concentrations below 20 µg m^{-3} throughout the day. The evening
357 peak was less distinct (Figure 6). On weekends, NO_2 concentrations remained below 20 µg
358 m^{-3} throughout the day with no distinct morning peak (Figure 6). Ozone followed an opposite
359 diurnal cycle to NO_2 with a maximum at site E and a minimum at sites A, C and G (see SI),
360 consistent with titration of marine background O_3 against vehicle-emitted NO (Weissert et al.,
361 2017).



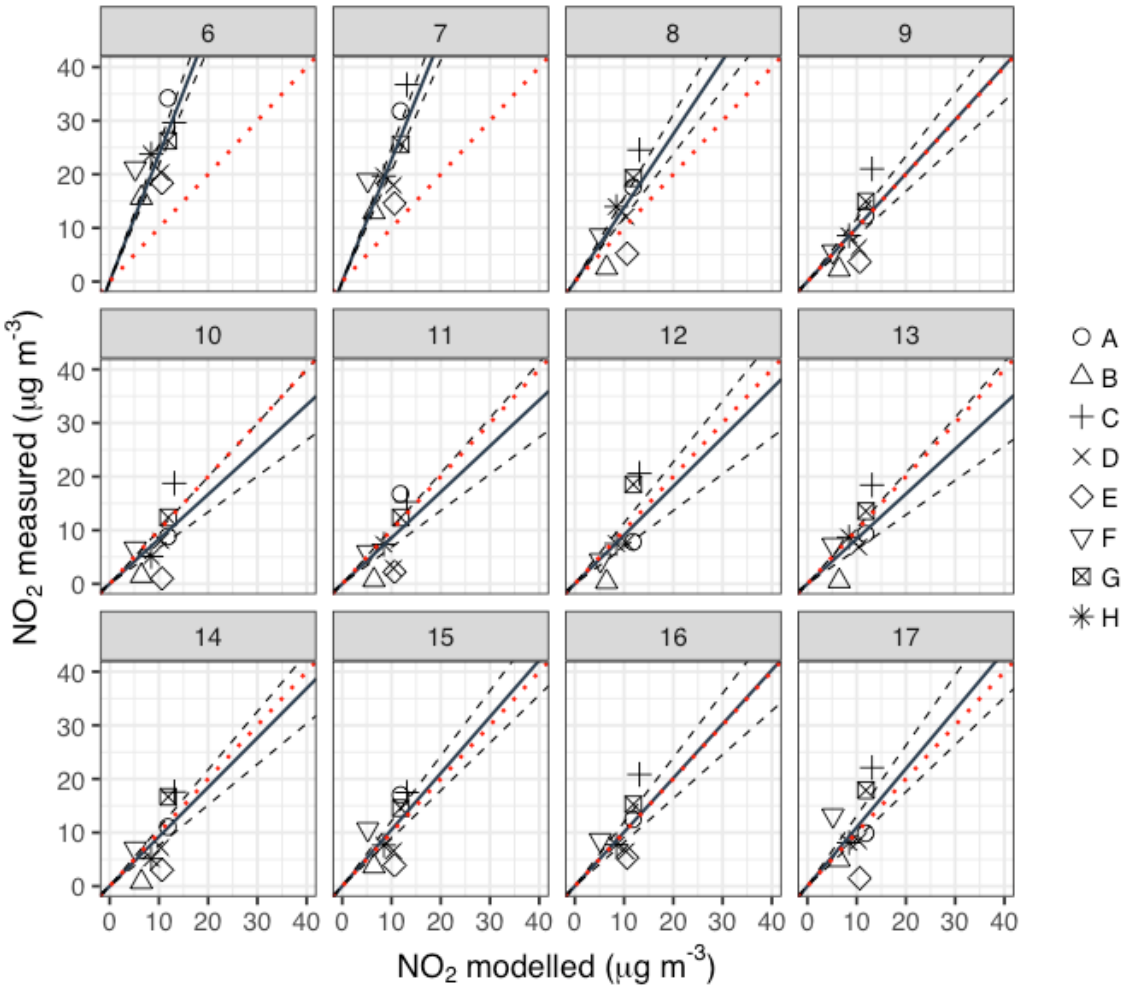
362

363 Figure 6. Comparison of the diurnal variability across sites of the a) mean and b) standard
364 deviation of NO_2 concentrations between weekday and weekends.

365

366 **4.2 Time-dependent LUR model**

367 Figure 7 gives example correlations according to equation 2 of the hourly-averaged sensor
368 measurements with the microscale LUR model, applied at the eight sites. It indicates that,
369 while the static linear model is reasonable (Adj. R^2 = 0.70 – 0.96 for the example shown in
370 Figure 7) there are some significant site-specific, time-dependent deviations.



371

372 Figure 7. Linear regression of hourly-averaged sensor results from one day against the static
373 LUR model applied at the eight instrument sites. The black solid line is the slope from
374 equation 2 and the black dashed lines are ± 1 standard error of the slope. The dotted red line
375 is the 1:1 line. The panels are different hours of the day.

376

377 Figure 8 shows the hourly-averaged sensor data plotted against the hourly-averaged modelled
378 NO_2 concentrations, $\bar{c}_{i,l}$ (equation 3), for each site over the entire deployment period. The
379 model in general captured the time- and space-variation in NO_2.

380

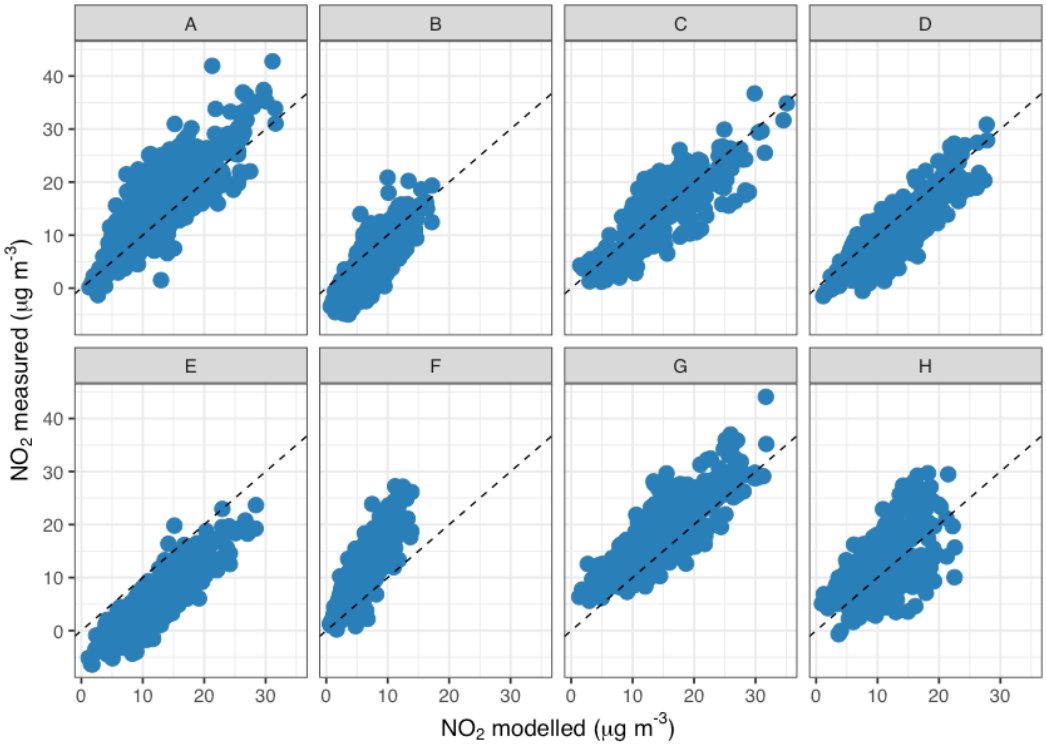
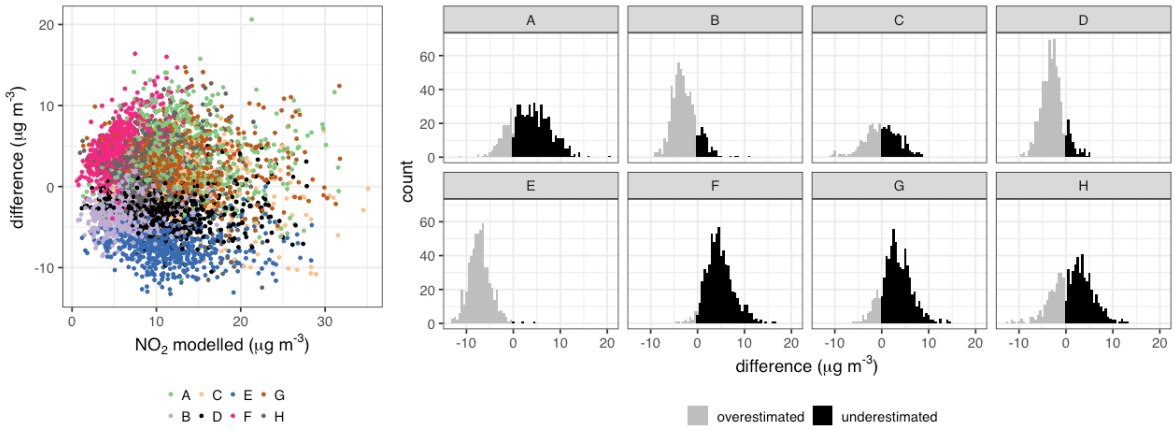


Figure 8. Hourly averaged NO_2 concentrations measured at the eight sites against the modelled NO_2 concentrations ($\bar{c}_{i,l}$, equation 3). The dashed line is the 1:1 line.

Figure 9 shows the difference between the modelled and measured concentrations, $\varepsilon_{l,k}$ (equation 3), coded by site, for the whole period of the study, plotted against the hourly-averaged modelled data and also shows the distribution of the difference at each site. Figure 9 indicates that there is no concentration-dependent bias. Figure 10 shows the mean diurnal difference averaged over the entire deployment period, between hourly-averaged measured and modelled NO_2 concentrations, evaluated for each site. These variations are small, but the errors are also correspondingly reduced by the averaging over 30 days: estimated standard deviation ~ 1 µg m^{-3}. Figures 9 and 10 show that there are site- and time-specific effects that are not captured by the model and that are not well-correlated with the modelled concentrations.



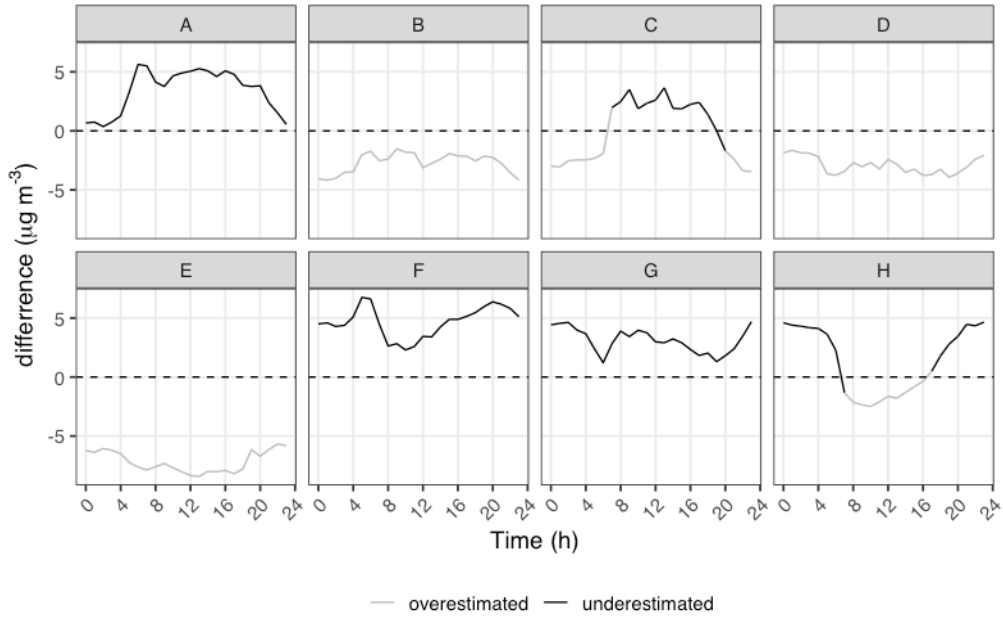
396

397 Figure 9. Variation of the difference between modelled and measured hourly-average NO_2

398 concentrations ($\varepsilon_{l,k}$, equation 3) with the modelled hourly-averaged concentration for the eight

399 instrument sites, and the distribution of this difference at the different sites.

400



401

402 Figure 10. The mean diurnal variability of $\varepsilon_{l,k}$ at the instrument sites, *k*. The colours show

403 where the model over- or underestimated NO_2 concentrations.

404

405 **5. Discussion**

406 Despite the low pollutant concentrations, the low-cost instruments successfully captured
407 significant differences between the different sites. The model, fusing the low-cost sensor
408 observations with the microscale LUR model, captured the time and space variation
409 reasonably well. The distribution of the difference between the modelled and measured NO_2
410 concentrations shown in Figure 10 shows small offsets for different instruments, that are
411 within the error estimates developed in the 'methods' section. Superimposed on these offsets,
412 a clear diurnal variation can be seen at some sites and not at others, of the difference between
413 measurement and model, averaged over the whole deployment (Figure 10). Although the
414 number of sites in the study was small, a tentative assessment of the site- and time-specific
415 effects could be made, relating the observations to specific urban design features. Table 1
416 shows the dominant parameters for the LUR model at each instrument site, and also notes the
417 side of the road, the presence of traffic lights and the direction with respect to the dominant
418 direction of traffic, which determines whether traffic is stationary during the rush hour or
419 accelerating on the instrument side of the road. The highest concentrations, which also had a
420 notable diurnal variation, were observed at a site under an awning directly next to a bus stop,
421 (site A and C, respectively). Sites close to bus stops but not close to traffic lights (D, E, F)
422 did not show a significant diurnal variation. Site H was unusual, showing a diurnal variation
423 with the lowest concentrations during the day.
424
425 LUR modelling studies undertaken at the local (Miskell et al., 2015) and microscale (Weissert
426 et al., 2018) in Auckland have identified the proximity and number of bus stops as key
427 predictors for NO_2 concentrations. The importance of bus stops within a short distance within
428 the LUR model emphasises the significance of diesel buses as the major source of NO_x
429 emissions. A sizable share of NO_x emissions from diesel vehicles is associated with uphill
430 driving, acceleration on a ramp, or positive accelerations from a standstill (Franco, 2014; Gis,

2017). For turbocharged diesel engines, NO emissions increase strongly with relative positive acceleration, defined as the integral of the product of instantaneous speed and instantaneous positive acceleration over a defined section of a driving schedule, and are also correlated with gear changes during the acceleration (Giakoumis, 2018; Giakoumis and Zachiotis, 2018). The occurrence of events such as these would vary across the day, and would depend on traffic lights, the dominant direction of traffic with respect to the lights, traffic density and the interaction of traffic lights with one another (reflected in the average transit time). A diurnal variation might thus be expected, associated with traffic lights. Sites D, E, F were some distance from traffic lights, which could account for the lack of diurnal variation of the difference between measurement and model at these sites. Site A, which recorded the highest concentrations, is upstream from a traffic light and on the west side of the road, where morning traffic is higher, and under an awning. It is directly by a bus stop. Further, it is on a slight hill, where traffic is accelerating away from the traffic light, and exit from the bus stop is impeded by parked cars. Buses at site A arrive around every 5 minutes throughout the day, likely explaining the continuously high NO_2 concentrations at this site. Whilst there is no bus stop very close to site C, this site is upstream from a traffic light, just after where a bus-only near-side lane terminates at a lane for cars turning left at the traffic light (driving is on the left in New Zealand). Buses, which pass this site around every 5 minutes throughout the day, are thus stopped near the site by the traffic waiting to turn (see SI for site photos). Whilst measured concentrations at Site H were low, this site was unusual in showing a diurnal variation with overestimated concentrations during the morning and underestimated concentrations in the evening. There are a number of fast-food outlets near this site, which are busiest at night, which may explain the observed variation. Although the offsets for the different instruments were within the error estimates for the model, some tentative deductions can be made in relation to the results from sites E, F and G. Site F was not directly next to a

bus stop but it was located between the entry and exit of a large car park hence where vehicles would be accelerating, possibly explaining the higher NO_2 concentrations measured by the instruments compared to the model. Similar to site A, site G was directly downstream from a traffic light, likely explaining the higher measured NO_2 concentrations. Site E was close to a bus stop and under an awning, where higher NO_2 concentrations would be expected, yet measured NO_2 concentrations were low. However, there is a potential effect of the siting: the instrument was against the wall directly under the awning, where air circulation may be limited. Previous studies have shown that pollutant concentrations can be significantly lower at the building side than the kerbside (Moodley et al., 2011) and that concentrations decrease with increasing height (Vardoulakis et al., 2011), both effects possibly explaining the lower NO_2 concentrations recorded by the instrument. The diffusion tubes, which were used to develop the LUR model, on the other hand were directly on the kerbside and at a lower height (2 m).

While NO_2 concentrations in Auckland are relatively low compared to other cities, our results indicate that the largest contribution of high NO_2 concentrations is likely related to buses, which are almost exclusively diesel operated in Auckland, and start-stop traffic at traffic lights. Considering that regular commuters wait at bus stops twice a day, five times per week, which accumulates to around 60 min per week, and wait at intersections to cross the road, this can contribute considerably to an individual's exposure (Velasco and Tan, 2016).

Table 1. Dominant parameters for the LUR model (bold parameters were significant in the LUR model development) and additional parameters that seem to be relevant at the instrument locations. Values are the characteristics of each of these parameters at each instrument site.

Site	**Awning**	**Nr. of bus stops**	Traffic light within 100 m	Distance to bus stop (m)	Side of Street	Dominant direction to traffic light

		within 100 m				
A	Y	3	Y	3	W	Downstream
B	N	3	Y	30	E	Upstream
C	Y	4	Y	70	W	Upstream
D	Y	2	N	18	W	-
E	Y	2	N	10	E	-
F	N	2	N	15	E	-
G	Y	3	Y	61	W	Downstream
H	Y	1	N	37	W	-

6. Conclusion

We have presented a novel approach to combine a network of low-cost air quality instruments with microscale LUR models to interpolate and map NO_2 concentrations at 50 m spatial scale and hourly time-scale. With attention to detail, specifically to the electrochemical NO_2 sensor noise, and to reliable correction for the ozone interference, low-cost NO_2 instruments give data that are sufficiently accurate to capture the important spatio-temporal variations of concentration of this pollutant in an urban environment, even if the concentrations are low. The study has also shown how analysis of differences between the measured and modelled concentrations can reveal specific urban design features that are not necessarily well-captured by simple linear LUR models but which might contribute disproportionately to population exposure. Further research is needed to assess site and time specific effects across different seasons and weather conditions. However, to date long-term NO_2 data from dense low-cost instrument networks are rare largely due to sensor drift and the lack of efficient calibration procedures.

The ability to map air pollutants across a grid for any given day and hour at a high spatial resolution offers new opportunities for exposure assessments that take account of people moving through microenvironments at different times of the day/week/year. For example, it allows comparing the air pollution exposure along different routes to a supermarket, school

or restaurant at different hours of the day, different days of the week or month. The method could be used as a tool to identify pollution hot spots at different times of the day, which could help urban planning.

This study has indicated the importance of site-specific interaction effects between land use variables, that are not well handled by simple spatial linear models. Models which incorporate both space- and time- dependent variables, and which can be constructed through a measurement programme using simple hand-held sensors (Miskell et al., 2018b) offer a possible way forward. Another approach may be the use of geographically weighted regression models, that account for non-stationary spatial effects (Song et al., 2019). Limitations are also imposed by siting requirement for the low-cost instruments, chiefly access to power. Specifically, there is the need to ensure that the sampled air is representative of the main land use. If the network is large enough, the LUR model may be built directly using the output from the low-cost air quality instruments. This may also allow investigation of space- and time-dependent effects and model adjustments in more detail.

Acknowledgments

We thank Gwynn Evans for his help during instrument placement and the shop owners for generously providing permission to install the instruments. This work was financially supported by the NZ Ministry of Business, Innovation and Employment, contract UOAX1413.

References

Air Quality Sensor Performance Evaluation Center, 2018. Field evaluation Aeroqual AQY (v0.5). Retrieved from http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/aeroqual-aqy-v0-5---field-evaluation.pdf?sfvrsn=4.

Aliwell, S.R., Halsall, J.F., Pratt, K.F., O'Sullivan, J., Jones, R.L., Cox, R.A., Utembe, S.R.,

525 Hansford, G.M., Williams, D.E., 2001. Ozone sensors based on WO_3: A model for
526 sensor drift and a measurement correction method. Measurement Science and
527 Technology 12, 684 - 690.
528 Auckland Transport, 2017. Dominion road upgrade.
529 Baron, R., Saffell, J.R., 2017. Amperometric gas sensors as a low cost emerging technology
530 platfrom for air quality monitoring applications: A review. ACS Sensors 2, 1553-
531 1566.
532 Bart, M., Williams, D.E., Ainslie, B., McKendry, I., Salmond, J., Grange, S.K., Alavi-
533 Shoshtari, M., Steyn, D., Henshaw, G.S., 2014. High density ozone monitoring using
534 gas sensitive semi-conductor sensors in the lower Fraser valley, British Columbia.
535 Environ Sci Technol 48, 3970-3977. 10.1021/es404610t
536 Cordero, J.M., Borge, R., Narros, A., 2018. Using statistical methods to carry out in field
537 calibrations of low cost air quality sensors. Sensors and Actuators B: Chemical 267,
538 245-254. 10.1016/j.snb.2018.04.021
539 Cordioli, M., Pironi, C., De Munari, E., Marmiroli, N., Lauriola, P., Ranzi, A., 2017.
540 Combining land use regression models and fixed site monitoring to reconstruct
541 spatiotemporal variability of NO_2 concentrations over a wide geographical area. Sci
542 Total Environ 574, 1075-1084. 10.1016/j.scitotenv.2016.09.089
543 Cross, E.S., Williams, L.R., Lewis, D.K., Magoon, G.R., Onasch, T.B., Kaminsky, M.L.,
544 Worsnop, D.R., Jayne, J.T., 2017. Use of electrochemical sensors for measurement
545 of air pollution: Correcting interference response and validating measurements.
546 Atmospheric Measurement Techniques 10, 3575-3588. 10.5194/amt-10-3575-2017
547 Deville Cavellin, L., Weichenthal, S., Tack, R., Ragettli, M.S., Smargiassi, A., Hatzopoulou,
548 M., 2016. Investigating the use of portable air pollution sensors to capture the spatial
549 variability of traffic-related air pollution. Environ Sci Technol 50, 313-320.

10.1021/acs.est.5b04235

Feinberg, S., Williams, R., Hagler, G.S.W., Rickard, J., Brown, R., Garver, D., Harshfield, G., Stauffer, P., Mattson, E., Judge, R., Garvey, S., 2018. Long-term evaluation of air sensor technology under ambient conditions in Denver, Colorado. Atmospheric Measurement Techniques 11, 4605-4615. 10.5194/amt-11-4605-2018

Franco, V., Sánchez, F. P., German, J., Mock, P., 2014. Real-world exhaust emissions from modern diesel cars - a meta-analysis of pems emissions data from EU (Euro 6) and US (tier 2 bin 5/ulev ii) diesel passenger cars. Part 1: Aggregated results. International Council on Clean Transportation Europe, Neue Promenade 6, 10178 Berlin.

Giakoumis, E.G., Triantafillou, G., 2018. Analysis of the effect of vehicle, driving and road parameters on the transient performance and emissions of a turbocharged truck. Energies 11, 295. 10.3390/en11020295

Giakoumis, E.G., Zachiotis, A.T., 2018. A comprehensive comparative investigation of a heavy-duty vehicle's performance, consumption and emissions during eight driving cycles. International Journal of Ambient Energy, 1-17. 10.1080/01430750.2018.1525578

Gis, M., 2017. Comparative studies exhaust emissions of the euro VI buses with diesel engine and spark-ignition engine cng fuelled in real traffic conditions. VII International Congress on Combustion Engines,. MATEC Web of Conferences.

Hansford, G.M., Freshwater, R.A., Bosch, R.A., Cox, R.A., Jones, R.L., Pratt, K.F., Williams, D.E., 2005. A low cost instrument based on a solid state sensor for balloon-borne atmospheric O_3 profile sounding. J Environ Monit 7, 158-162. 10.1039/b412184h

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D.,

2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmospheric Environment 42, 7561 - 7578.

Hossain, M., Saffell, J.R., Baron, R., 2016. Differentiating NO_2 and O_3 at low cost air quality amperometric gas sensors. ACS Sensors 1, 1291 - 1294.

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 15, 185-204. 10.1038/sj.jea.7500388

Johnson, K., MacNeill, M., Grgicak-Mannion, A., Nethery, E., Xu, X., Dales, R., Rasmussen, P., Wheeler, A.J., 2013. Development of temporally refined land-use regression models predicting daily household-level air pollution in a panel study of lung function among asthmastic children. Journal of Exposure Science and Environmental Epidemiology 23, 259 - 267.

Lewis, A.C., Lee, J.D., Edwards, P.M., Shaw, M.D., Evans, M.J., Moller, S.J., Smith, K.R., Buckley, J.W., Ellis, M., Gillot, S.R., White, A., 2016. Evaluating the performance of low cost chemical sensors for air pollution research. Faraday Discussions 189, 85 - 103.

Lin, C., Masey, N., Wu, H., Jackson, M., Carruthers, D., Reis, S., Doherty, R., Beverland, I., Heal, M., 2017. Practical field calibration of portable monitors for mobile measurements of multiple air pollutants. Atmosphere 8, 231. 10.3390/atmos8120231

Masiol, M., Zikova, N., Chalupa, D.C., Rich, D.Q., Ferro, A.R., Hopke, P.K., 2018. Hourly land-use regression models based on low-cost pm monitor data. Environ Res 167, 7-14. 10.1016/j.envres.2018.06.052

Mead, M.I., Popoola, O.A.M., Stewart, G.B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J.J., McLeod, M.W., Hodgson, T.F., Dicks, J., Lewis, A., Cohen, J., Baron,

R., Saffell, J.R., Jones, R.L., 2013. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. Atmospheric Environment 70, 186-203. 10.1016/j.atmosenv.2012.11.060

Miskell, G., Salmond, J., Alavi-Shoshtari, M., Bart, M., Ainslie, B., Grange, S., McKendry, I.G., Henshaw, G.S., Williams, D.E., 2016. Data verification tools for minimizing management costs of dense air-quality monitoring networks. Environ Sci Technol 50, 835-846. 10.1021/acs.est.5b04421

Miskell, G., Salmond, J., Longley, I., Dirks, K.N., 2015. A novel approach in quantifying the effect of urban design features on local-scale air pollution in central urban areas. Environ Sci Technol 49, 9004-9011. 10.1021/acs.est.5b00476

Miskell, G., Salmond, J.A., Williams, D.E., 2018a. Solution to the problem of calibration of low-cost air quality measurement sensors in networks. ACS Sensors 3, 832-843. 10.1021/acssensors.8b00074

Miskell, G., Salmond, J.A., Williams, D.E., 2018b. Use of a handheld low-cost sensor to explore the effect of urban design features on local-scale spatial and temporal air quality variability. Science of The Total Environment 619-620, 480-490. 10.1016/j.scitotenv.2017.11.024

Moodley, K.G., Singh, S., Govender, S., 2011. Passive monitoring of nitrogen dioxide in urban air: A case study of Durban metropolis, South Africa. J Environ Manage 92, 2145-2150. 10.1016/j.jenvman.2011.03.040

Mukerjee, S., Smith, L., Neas, L., Norris, G., 2012. Evaluation of land use regression models for nitrogen dioxide and benzene in four us cities. Scientific World Journal 2012, 865150. 10.1100/2012/865150

Munir, S., Mayfield, M., Coca, D., Jubb, S.A., Osammor, O., 2019. Analysing the performance of low-cost air quality sensors, their drivers, relative benefits and

625 calibration in cities - a case study in Sheffield. Environmental Monitoring and
626 Assessment 191.
627 Nethery, E., Teschke, K., Brauer, M., 2008. Predicting personal exposure of pregnant
628 women to traffic-related air pollutants. Sci Total Environ 395, 11-22.
629 10.1016/j.scitotenv.2008.01.047
630 Pang, X., Shaw, M.D., Gillot, S., Lewis, A.C., 2018. The impacts of water vapour and co-
631 pollutants on the performance of electrochemical gas sensors used for air quality
632 monitoring. Sensors and Actuators B: Chemical 266, 674-684.
633 10.1016/j.snb.2018.03.144
634 Pang, X., Shaw, M.D., Lewis, A.C., Carpenter, L.J., Batchellier, T., 2017. Electrochemical
635 ozone sensors: A miniaturised alternative for ozone measurements in laboratory
636 experiments and air-quality monitoring. Sensors and Actuators B: Chemical 240, 829-
637 837. 10.1016/j.snb.2016.09.020
638 Popoola, O.A.M., Carruthers, D., Lad, C., Bright, V.B., Mead, M.I., Stettler, M.E.J., Saffell,
639 J.R., Jones, R.L., 2018. Use of networks of low cost air quality sensors to quantify air
640 quality in urban settings. Atmospheric Environment 194, 58-70.
641 10.1016/j.atmosenv.2018.09.030
642 Senaratne, I., Shooter, D., 2004. Elemental composition in source identification of brown
643 haze in Auckland, New Zealand. Atmospheric Environment 38, 3049-3059.
644 10.1016/j.atmosenv.2004.02.046
645 Slama, R., Morgenstern, V., Cyrys, J., Zutavern, A., Herbarh, O., Wichmann, H.E.,
646 Heinrich, J., LISA Study Group, 2007. Traffic-related atmospheric pollutants levels
647 during pregnancy and offspring's term birth weight: A study relying on a land-use
648 regression exposure model. Environmental Health Perspectives 115, 1283 - 1292.
649 Son, Y., Osornio-Vargas, Á.R., O'Neill, M.S., Hystad, P., Texcalac-Sangrador, J.L.,

650 Ohman-Strickland, P., Meng, Q., Schwander, S., 2018. Land use regression models to
651 assess air pollution exposure in Mexico city using finer spatial and temporal input
652 parameters. Science of The Total Environment 639, 40-48.
653 10.1016/j.scitotenv.2018.05.144
654 Song, W., Jia, H., Li, Z., Tang, D., Wang, C., 2019. Detecting urban land-use configuration
655 effects on NO_2 and no variations using geographically weighted land use regression.
656 Atmospheric Environment 197, 166-176. 10.1016/j.atmosenv.2018.10.031
657 Statistics New Zealand, 2013. 2013 census usually resident population counts.
658 Utembe, S.R., Hansford, G.M., Sanderson, M.G., Freshwater, R.A., Pratt, K.F.E., Williams,
659 D.E., Cox, R.A., Jones, R.L., 2006. An ozone monitoring instrument based on the
660 tungsten trioxide (WO_3) semiconductor. Sensors and Actuators B: Chemical 114, 507-
661 512. 10.1016/j.snb.2005.04.049
662 Vardoulakis, S., Solazzo, E., Lumbreras, J., 2011. Intra-urban and street scale variability of
663 btex, NO_2 and O_3 in Birmingham, UK: Implications for exposure assessment.
664 Atmospheric Environment 45, 5069-5078. 10.1016/j.atmosenv.2011.06.038
665 Velasco, E., Tan, S.H., 2016. Particles exposure while sitting at bus stops of hot and humid
666 Singapore. Atmospheric Environment 142, 251-263.
667 10.1016/j.atmosenv.2016.07.054
668 Weissert, L.F., Salmond, J.A., Miskell, G., Alavi-Shoshtari, M., Grange, S.K., Henshaw,
669 G.S., Williams, D.E., 2017. Use of a dense monitoring network of low-cost
670 instruments to observe local changes in the diurnal ozone cycles as marine air passes
671 over a geographically isolated urban centre. Sci Total Environ 575, 67-78.
672 10.1016/j.scitotenv.2016.09.229
673 Weissert, L.F., Salmond, J.A., Miskell, G., Alavi-Shoshtari, M., Williams, D.E., 2018.
674 Development of a microscale land use regression model for predicting NO_2

concentrations at a heavy trafficked suburban area in Auckland, NZ. Sci Total Environ 619-620, 112-119. 10.1016/j.scitotenv.2017.11.028

Williams, D.E., Aliwell, S.R., Pratt, K.F., Caruana, D.J., Jones, R.L., Cox, R.A., Hansford, G.M., Halsall, J., 2002. Modelling the response of a tungsten oxide semiconductor as a gas sensor for the measurement of ozone. Measurement Science and Technology 13, 923-931.

Williams, D.E., Henshaw, G.S., Bart, M., Laing, G., Wagner, J., Naisbitt, S., Salmond, J.A., 2013. Validation of low-cost ozone measurement instruments suitable for use in an air-quality monitoring network. Measurement Science and Technology 24, 065803. 10.1088/0957-0233/24/6/065803

Xie, S., Sridhar, S., Metcalfe, J., 2014. Auckland air emissions inventory, Auckland Council technical report, TR2014/015. Auckland Council Auckland, New Zealand.

Yeganeh, B., Hewson, M.G., Clifford, S., Tavassoli, A., Knibbs, L.D., Morawska, L., 2018. Estimating the spatiotemporal variation of NO_2 concentration using an adaptive neuro-fuzzy inference system. Environmental Modelling & Software 100, 222-235. 10.1016/j.envsoft.2017.11.031