Research Article

A Simple Tool to Identify Representative Wind Sites for Air Pollution Modelling Applications

M. A. Elangasinghe,^{1,2} K. N. Dirks,³ N. Singhal,¹ J. A. Salmond,⁴ I. Longley,⁵ and V. I. Dirks⁶

¹Faculty of Engineering, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

²Faculty of Engineering, University of Peradeniya, 20400 Peradeniya, Sri Lanka

³Faculty of Medical and Health Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

⁴Faculty of Science, University of Auckland, Private Bag 92019, Auckland, 1142, New Zealand

⁵National Institute of Water and Atmospheric Research (NIWA) Ltd., 41 Market Place, Viaduct Harbour, Auckland Central 1010, New Zealand

⁶44B Kelvin Road, Remuera, Auckland 1050, New Zealand

Correspondence should be addressed to M. A. Elangasinghe; mela005@aucklanduni.ac.nz

Received 20 March 2015; Accepted 13 August 2015

Academic Editor: George Kallos

Copyright © 2016 M. A. Elangasinghe et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigates the use of the Site-Optimized Semiempirical (SOSE) air pollution model to identify the surface wind measurement site characteristics that yield the best air pollution predictions for urban locations. It compares the modelling results from twelve meteorological sites with varying anemometer heights, located at different distances from the air pollution measurements and exhibiting different land use characteristics. The results show that the index of agreement (IA) between observed and predicted concentrations can be improved from 0.4 to 0.8 by using the most compared to the least representative wind data as input to the air pollution model. Although improvements can be achieved using wind data from a site closer to the air quality monitoring site, choosing the closest wind site does not necessarily yield the best results, especially if the meteorological station is located in a region of complex land use. In addition, both the height of the anemometer and the openness of the terrain surrounding the anemometer were found to be equally important in obtaining good model predictions. The simple SOSE model can therefore be used to complement regulatory meteorological guidelines by providing a quantitative assessment of wind site representativeness for air quality applications in complex urban environments.

1. Introduction

The success of a meteorological air pollution model depends to a large extent on the “representativeness” of its input data in describing the meteorological characteristics of the atmosphere which govern the dispersion of pollutants [1, 2]. For the purpose of assessing air quality, the World Meteorological Organization (WMO) recommends that wind measurements should be made at a height of ten metres or 1.5 times the mean height of roughness elements in densely built-up areas and that the measurements should be made in an unobstructed area, with an “unobstructed area” defined as “an area where the distance between the instrument and any obstruction is at

least ten times the height of that obstruction” [3–5]. Siting and exposure requirements of meteorological monitoring stations specifically for the purpose of air quality modelling are also documented by the US Environmental Protection Agency [6]. Colocated meteorological measurements (measurements made in close proximity to an air quality monitoring site and specifically for the purpose of studying air quality) are considered to be preferable to national weather station data for use in air quality modelling, providing that appropriate instrumentation is used and that suitable quality assurance procedures are followed [7]. In the absence of site-specific measurements, national weather station data are considered acceptable by US Environmental Protection Agency [7],

even though their prime focus is often on monitoring and predicting severe and adverse weather while the focus of air quality studies is on mild weather conditions [8].

In accordance with the above guidelines, most urban air pollution modelling analyses [9–12] use meteorological data that are either colocated or located as close as possible to the air pollution monitor in question. When colocated wind measurements are not available, meteorological data from weather stations located outside of the urban areas (e.g., at an airport site) are sometimes used [13–16]. Considering the comprehensive and long-term meteorological data availability of weather stations, some studies have resorted to using wind data from stations located 20–50 km away [1, 15, 16]. Practical constraints over the choice of sites, monitoring expense, and the desirability of colocating meteorological measurements with air quality instruments (where site selection is often constrained by other variables) mean that, in many cases, compromises have to be made in terms of meeting the siting requirements. Moreover, since topography, land use, and roughness length are inherently variable in urban areas, it may not be possible to identify any sites that comply with all of the requirements, as specified by the USEPA or WMO [2].

At the present time, the choice of the most suitable meteorological site for a particular air pollution modelling application remains somewhat arbitrary and based on limited understanding of the impact of the various meteorological site characteristics in air quality model performance. Moreover, there is no generally accepted analytical or statistical technique specified to determine the representativeness of meteorological data or monitoring sites for the purpose of air quality modelling [6]. A few studies have been carried out which evaluate the representativeness of wind data on air quality model performance [11, 17]. One study carried out in the city of Florence compared the performance accuracy of an air quality model between wind data from meteorological observations and numerical weather prediction [11]. A similar study has been carried out in Italy comparing the performance of a model using observed urban meteorological data over meteorological data provided by the CALMET preprocessor [17]. However, as far as the authors are aware, there are no tools available to evaluate model performance using different *observational* data in relation to the “representativeness” of the meteorological sites on which the model outputs are based. The purpose of this paper is therefore to use simple Site-Optimized Semiempirical (SOSE) model as a tool for identifying the relative importance of meteorological site characteristics (such as proximity of meteorological monitoring to an air quality monitoring site, the height of the anemometer, and whether or not the site is located in an “open” area) in relation to air pollution model performance.

2. Materials and Methods

2.1. Site Selection. The study was carried out in Auckland, the largest city in New Zealand, located along a narrow isthmus consisting of a complex coastline. The topography is

FIGURE 1: Map of monitoring locations used in this study. (a) Map showing twelve wind sites; (b) enlarged view of the three air pollution measurement sites with on-site anemometers for wind measurements.

low-lying undulating terrain. Due to the limited influence of industrial emissions, vehicular emissions and home heating are the major sources of air pollution. The key pollutants that impact air quality in the Auckland region therefore are CO , NO_x , and PM_{10} [18].

Air pollution concentrations used for this study are from the ROADSIDE field campaign [19] that covers a period of four months at a site focusing on Auckland’s Southern Motorway, between 2 April 2010 and 1 August 2010. The site is located in Otahuhu East, a suburban residential neighbourhood in Auckland with generally flat terrain which is bisected by Auckland’s Southern Motorway (annual average daily traffic volumes of ~120,000). Air quality data and colocated meteorological data were collected from three monitoring sites located within 440 m of each other in close proximity to the Southern Motorway (see Figure 1). The locations of each of the air quality monitoring stations are shown in Figure 1(b). Luke Street is located in the west of the motorway set back by 240 m in open grassy terrain relatively free of any building or natural structures and so, to a large extent, complies with the WMO siting requirements. The 25 Deas Place site is also located 140 m from the motorway but to the east. It is located within a suburban neighbourhood consisting of single-story residential housing with modest amounts of vegetation (small trees and shrubs). Deas Place Reserve site is located in very close proximity (approximately 5 m from the motorway) to a slip road that is an exit of

TABLE 1: Air pollution and meteorological data.

Site number	Site name	Abbreviation	Monitoring data	H above MSL (m)	H_A^b (m)	Distance (km)	Costal/inland	Nature of the site
1	Luke Street ^a	LS	Wind, CO, NO, NO ₂ PM ₁₀	15	10	—	Inland	Suburban, open
2	Deas Place Reserve ^a	DPR	Wind, CO, NO, NO ₂ PM ₁₀	4	6	—	Inland	Next to motorway, suburban, covered
3	25 Deas Place ^a	25DP	Wind, CO, NO, NO ₂ PM ₁₀	0.2	8	—	Inland	Suburban, residential, covered
4	Henderson	H	Wind only	35	10	21	Inland	Suburban, covered
5	Kyberpass	K	Wind only	81	12	10	Inland	Urban, Covered
6	Mangere	M	Wind only	5	10	7	Coastal	Rural, open
7	Music Point	MP	Wind only	18	10	10	Coastal	Rural, covered
8	Onehunga	O	Wind only	5	10	5	Coastal	Suburban, open
9	Pakuranga	PA	Wind only	15	6	5	Inland	Urban, covered
10	Penrose	PE	Wind only	30	6	5	Inland	Suburban, open
11	Pukekohe	PU	Wind only	88	10	30	Inland	Rural, open
12	Wiri	W	Wind only	18	10	7	Inland	Suburban, open

^aSites with colocated wind data and air pollution data, ^b H_A : anemometer height.

the motorway. The air pollution data consists of 10-minute averages of CO, NO₂, NO_x, and PM₁₀ measured at 5 m above ground level. The colocated meteorological towers are fixed at a height of 10 m at Luke Street, 8 m at 25 Deas Place, and 6 m at Deas Place Reserve.

The meteorological dataset was supplemented by data from nine other meteorological monitoring sites across the region (see Figure 1(a)), downloaded from the National Climate Database maintained by NIWA (<http://cliflo.niwa.co.nz/>). Each of the twelve meteorological monitoring sites was different in terms of their site characteristics, including their measurement heights, their distance to the air quality monitoring sites, and the land use characteristics of the surrounding area (building height, density, etc.). The site-specific characteristics of the twelve sites considered for the study are given in Table 1. The abbreviations used in Table 1 for the twelve sites are used here onwards.

2.2. The Model. The air pollution model chosen for this study is a simple Site-Optimized Semiempirical model (SOSE) [9, 10]. This model was developed and tested in New Zealand [9, 10] and shown to be effective in predicting ambient concentrations of a range of pollutants associated with road traffic [20]. It has been found to be useful in practical applications such as interpolating for missing data and looking at “what if” scenarios associated with changes in traffic patterns and surface meteorology [9, 10, 20, 21]. It has also been shown to be effective in conditions of complex terrain, as found in the Aosta Valley of Italy [20]. An advantage of this model is that it can be trained exclusively using a set of wind and concentration data and it can be expected to perform well if the wind data used as input are representative of

the area in terms of its dispersion characteristics. Carbon monoxide (CO) and oxides of nitrogen (NO_x) are chosen as the pollutants of interest as they are strongly associated with road traffic, as discussed above.

SOSE assumes the concentration, C , is inversely related to the wind speed (u) (as with the box model) but with a wind speed offset u_0 (ms⁻¹) included to avoid severe over-predictions in very light wind speed conditions, as suggested by Chock [22]. The model becomes

$$C = \frac{Q}{\Delta z (u + u_0)} + C_B, \quad (1)$$

where Q is the emission rate, Δz is the mixing height (or box height), and C_B is the background concentration of the pollutant. This equation is separately applied when the receptor is placed windward and leeward of the road. So for leeward (downwind) conditions, the equation becomes

$$C = \frac{Q_l}{\Delta z (u + u_0)} + C_l. \quad (2)$$

The emission term, Q_l , for leeward conditions incorporates both emissions from the road adjacent to the monitor and emissions from other roads in the vicinity and C_l is the background concentration for leeward condition.

For windward (upwind) conditions, the model becomes

$$C = \frac{Q_w}{\Delta z (u + u_0)} + C_w. \quad (3)$$

The emission term, Q_w , is the emission component from other roads in the vicinity and C_w is the background concentration for windward condition.

TABLE 2: Model evaluation statistics and their definitions.

Name	Symbol	Definition
Normalized root mean squared error	NRMSE	$\frac{\sqrt{(1/N) \sum_{i=1}^N (P_i - O_i)^2}}{O_{\max} - O_{\min}}$
Index of agreement	IA	$1 - \frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (P_i - \bar{O} + O_i - \bar{O})^2}$
Correlation coefficient	COR	$\frac{(1/N) \sum_{i=1}^N (C_o - \bar{C}_o)(C_p - \bar{C}_p)}{\sigma_{C_p} \sigma_{C_o}}$
Fractional bias	FB	$2 \frac{(\bar{P} - \bar{O})}{\bar{P} + \bar{O}}$
Factor of two	FAC2	Fraction of data for which $0.5 \leq P_i/O_i \leq 2$

Note. P_i is the i th predicted value, O_i is the i th observed value, $O_{\max}$ is the maximum observed value, $O_{\min}$ is the minimum observed value, N is the number of observed and predicted pairs, $\bar{P}$ is the mean predicted, $\bar{O}$ is the mean observed, σ_{C_p} is the standard deviation over the observed data set, and σ_{C_o} is the standard deviation over the predicted data set.

With the data sorted by time of day, linear regressions of C (mgm^{-3}) on $(u + u_0)^{-1}$ are performed and giving values of the regression coefficients $Q_l \Delta z^{-1}$, C_l for leeward conditions, and $Q_w \Delta z^{-1}$, C_w for windward conditions [9] for each time of day. If the daily distribution of emissions is different between weekdays and weekend days, the dataset may be partitioned so that regression parameters are obtained for weekday and weekend days separately.

Based on the three-month ROADSIDE dataset, the optimized model parameters, namely, $Q_l \Delta z^{-1}$, C_l , $Q_w \Delta z^{-1}$, and C_w , were calculated for weekdays and weekend days for each 10-minute period throughout the day with the optimum parameters being constrained to avoid negative concentration predictions.

2.3. The Procedure. The concentrations of CO and NO_x at the three air pollution monitoring locations were modelled using wind field observations from each of the twelve wind sites separately (three site-specific wind observational sites and nine from the Auckland climate network for the same time period). This resulted in 36 combinations of model results for each of the pollutants.

2.4. Model Evaluation Statistics. Model performance was evaluated using standard model evaluation statistics recommended in the literature for estimating the uncertainty in air quality model predictions [23, 24]. The statistics used were the normalized root mean squared error (NRMSE), the index of agreement (IA), the correlation coefficient (COR), the fractional bias (FB), and the fraction within a factor of two (FAC2). These statistics also have the advantage that they are dimensionless, allowing for easy comparison between pollutants. Their definitions are given in Table 2.

NRMSE and IA indicate the degree of agreement between observed and predicted time series data, and FB is a measure of agreement with the mean concentration; a positive FB will

result if the model is overpredicting the mean concentration and a negative FB will result if mean concentrations are underpredicted by the model. A perfect model performance will result in NRMSE and FB scores of zero and an IA of unity. FAC2 is a measure of the proportion of predictions within a factor two of the observed concentration.

3. Results and Discussion

3.1. Meteorological Data. Wind rose diagrams were constructed for the twelve wind sites for the period of April 1 to August 1, 2010. Some of these diagrams that show similar and different wind roses to that of the air pollution measurement site of interest are presented in Figure 2. The dominant wind flows for most of the sites during the observational period are from the south west and north east directions. Observations at LS, PA, and M were similar in terms of wind direction (Figures 2(a), 2(b), and 2(c)). However, relative weakening of intensities was observed at LS (moderate) and PA (weak). Based on the aerial view of the sites, the degree of surface cover around the site increases in the same order (M, LS, and PA). From this, we speculate that the weakening of winds is caused by increased surface roughness. The wind rose patterns at the Wiri (W) site, situated 7 km away, and Pukekohe (PU), the furthest site situated 30 km away, show strong north westerly wind components that are not observed at other sites (Figures 2(d) and 2(e)). At Onehunga (O), located 5 km away from the study site, wind components from the east and south-east were observed, components that were not found at any of the other closer sites (Figure 2(f)). All of the other wind roses showed significant differences from those of the LS, PA, and M sites (not shown here).

3.2. Air Quality Data. Figure 3(a) shows the normalized time variation of CO, NO_x, and wind speed at 25DP averaged for each hour of the day for the study period. Also shown

FIGURE 2: Wind rose diagrams for (a) Mangere, (b) Luke Street, (c) Pakuranga, (d) Pukekohe, (e) Wiri, and (f) Onehunga.

on the plots is the 95% confidence interval of the means calculated through bootstrap resampling. The distributions of CO and NO_x are consistent with peaks during the morning and evening rush hours with concentrations persisting into the early night. An inverse relationship between pollutant concentrations and wind speed is also well depicted from Figure 3(a).

Hourly average concentrations of CO and NO_x are calculated separately for the events when the receptor site 25DP is upwind or downwind with respect to the motorway and is presented in Figures 3(b) and 3(c), respectively. Both CO and NO_x concentration averages are higher when the site

is downwind relative to upwind of the motorway. This is consistent with the presence of a significant line source and highlights the need to treat upwind and downwind time periods separately when modelling pollutants in the presence of such a source.

3.3. SOSE Modeling Results

3.3.1. Model Performance. Examples of SOSE modeling results for the fourth week of the four-month campaign, using ten-minute averages, based on colocated wind data from 25DP, LS, and DPR for observed CO and NO_x concentrations

FIGURE 3: (a) Time variation of hourly average concentration of CO, NO_x, and wind speed at 25DP (normalized values are presented). (b) Upwind and downwind concentrations of CO at 25DP. (c) Upwind and downwind concentrations of NO_x at 25DP. Note. The band shows 95% confidence interval through bootstrap resampling.

at 25DP are presented in Figure 4. Similar results were obtained for the remaining weeks of observation. Model statistics for the four-month campaign are presented in Tables 3 and 4. Figures 4(a) and 4(b) illustrate the modeling results using wind measurements at LS and highlight the enhanced model performance achieved when meteorological data from this same site are used to train the model. Figures 4(c) and 4(d) show the results for 25DP using data from colocated instruments at 25DP, and Figures 4(e) and 4(f) illustrate the results using data from DPR. In each of these latter four cases, model performance is poor, indicating that

the meteorological data used were not representative of the domain.

Figure 5 presents the time series of observed and predicted concentrations for one week using wind data from five weather stations from across the rest of the Auckland region (the locations of the weather stations are marked in Figure 1) using ten-minute averages. Figure 5(a) provides an example of the modelling results achieved using LS winds for the same week for comparison. The weather station at Mangere (M) appears to be the most representative of the nine noncolocated sites considered in this study (Figure 5(b)).

FIGURE 4: Examples of one-week time series of observed and predicted ten-minute average CO and NO_x concentrations at 25DP using wind data from 25DP, LS, and DPR, respectively. The results suggest that the best model performance is achieved using data from LS (the “open” site).

Wind measurements at PA were able to pick up some of the variability but do not represent the peaks in concentration well (Figure 5(c)). Wind measurements at Onehunga (O) and Wiri (W) lead to overpredictions of the peak pollutant concentrations, while wind data from Pukekohe (PU) result in marked underpredictions of the peak concentrations (Figures 5(d), 5(e), and 5(f)).

Table 3 presents the model evaluation statistics for NO_x concentrations at 25DP using wind data from the twelve meteorological stations. The results show that the best performance (NRMSE = 0.08, IA = 0.86, COR = 0.78, FB = 0.03, and FAC2 = 0.6) is obtained using wind data from LS located approximately 300 m away from the air quality monitoring site 25DP. The second best performance is obtained using

TABLE 3: Statistical analysis of the predicted and measured ten-minute average NO_x concentration at 25DP from April to July 2010.

Statistical parameter	Meteorological datasets from onsite meteorological stations			Meteorological datasets from Auckland climate network (arranged according to the distance from air pollution site, from left to right)								
	25DP	DPR	LS	PA	O	PE	M	W	K	MP	H	PU
NRMSE (range: 0-1, ideal: 0)	0.18	0.17	0.08	0.09	0.25	0.13	0.09	0.17	0.1	0.13	0.14	0.14
IA (range: 0-1, ideal: 1)	0.69	0.65	0.86	0.81	0.54	0.69	0.82	0.64	0.72	0.56	0.61	0.46
COR (range: 0-1, ideal: 1)	0.61	0.51	0.78	0.72	0.48	0.52	0.73	0.49	0.62	0.38	0.4	0.28
FB (range: 0-1, ideal: 0)	0.4	0.27	0.03	0.03	0.48	0.17	0.02	0.21	0.02	-0.08	-0.25	-0.77
FAC2 (range: 0-1, ideal: 1)	0.53	0.56	0.60	0.60	0.52	0.50	0.59	0.52	0.52	0.48	0.50	0.46

TABLE 4: Statistical analysis of the predicted and measured ten-minute average CO concentration at 25DP from April to July 2010.

Statistical parameter	Meteorological datasets from onsite meteorological stations			Meteorological datasets from Auckland climate network (arranged according to the distance from air pollution site, from left to right)								
	25DP	DPR	LS	PA	O	PE	M	W	K	MP	H	PU
NRMSE (range: 0-1, ideal: 0)	0.1	0.1	0.06	0.06	0.13	0.07	0.06	0.1	0.07	0.08	0.08	0.09
IA (range: 0-1, ideal: 1)	0.67	0.67	0.82	0.8	0.59	0.66	0.78	0.64	0.7	0.51	0.6	0.44
COR (range: 0-1, ideal: 1)	0.53	0.51	0.74	0.7	0.48	0.52	0.69	0.47	0.6	0.38	0.42	0.32
FB (range: 0-1, ideal: 0)	0.39	0.25	0.04	0.05	0.47	0.08	0.03	0.23	0.02	-0.13	-0.33	-0.88
FAC2 (range: 0-1, ideal: 1)	0.35	0.37	0.39	0.38	0.36	0.33	0.37	0.34	0.33	0.29	0.25	0.20

wind data from M and PA, the two sites where similar wind patterns to those at LS are observed. The worst model performance (NRMSE = 0.14, IA = 0.46, COR = 0.28, FB = -0.77, and FAC2 = 0.46) was obtained using wind data from PU, the site located the furthest away and displaying a noticeably different wind rose pattern. The agreement between observed and predicted concentrations reduced by 46% when using wind data from PU site compared with data from LS. All of the other sites (W, MP, H, and O), which showed markedly different wind roses to that at LS, resulted in poor model performance. According to the FB values, MP, H, and PU tended to result in underprediction of concentrations compared to the mean. The effectiveness of model predictions (as shown by FAC2) also declined in a similar order from LS (best wind site) to PU (worst wind site).

The same statistical analysis was repeated, this time using CO as the pollutant. The results are presented in Table 4. In terms of the ranking of the sites for model performance, the results are similar to those found for NO_x . Wind data from LS were found to produce the best model results (NRMSE = 0.06, IA = 0.82, COR = 0.74, FB = 0.04, and FAC2 = 0.39), while the results using PU data were found to be the worst (NRMSE = 0.09, IA = 0.44, COR = 0.32, FB = -0.88, and FAC2 = 0.2). The agreement between observed and predicted concentrations reduced by 46% when using wind data from the PU site compared with LS. The fractional bias takes a large negative

value with PU winds, which is a result of the underpredicted events. FAC2 values reduced from the best wind site to the worst wind site showing a similar pattern to NO_x .

These results show a similar level of improvement to the model as those achieved by Giambini et al. [11] using the model SIRANE when they replaced numerically predicted wind data with observed wind data (COR improved from 0.69 to 0.73) and the improvement to ADMS-Urban model obtained by Righi et al. [17] when wind data from the centre of the study area are used rather than those from a slightly remote station (COR improved from 0.2 to 0.45).

3.3.2. Model Performance in relation to Site Characteristics

Distance to Wind Site. The effect of distance between the anemometer and the air pollution monitor was evaluated by plotting the IA values of modelled results as a function of the distance (Figure 6(a)). The linear relationship between distance to wind site and IA values showed a correlation coefficient of 0.6 ($p = 0.03$). The results suggest that the representativeness of wind data is reduced with the distance to the wind site. However, the weak correlation is a result of wind data from some closer sites such as 25DP, DPR (on site), and PE (5 km away) that have their anemometers located in obstructive environments (with trees and buildings in close

FIGURE 5: The SOSE model results in predicting one week of NO_x concentration levels at 25DP using wind data from different wind sites; (a) Luke Street (LS), (b) Mangere (M), (c) Pakuranga (PA), (d) Wiri (W), (e) Onehunga (O), and (f) Pukekohe (PU).

FIGURE 6: (a) The effect of distance to the wind site on model performance at 25DP. (b) The effect of anemometer height and openness on the model performance at 25DP with wind data from three on-site anemometers.

proximity) and were showing low IA values compared to other close sites.

Height of Anemometer and Openness of the Site. When the IA values of modelled results at 25DP using three on-site meteorological stations were analyzed (Figure 6(b)), LS, the most open site having an anemometer height of 10 m, gives the best model performance. Prediction results for 25DP site using its colocated wind data were worse than those with LS wind data since the anemometer at 25DP was at 8 m and in an area covered with buildings. DPR wind data also showed less representativeness as a result of low anemometer height (6 m) and turbulence from the surrounding environment. This suggests that if the air pollution monitor is located in an area where it is not possible to erect an anemometer at 10 m in open grounds, it would be preferable to choose a wind monitoring site away from the air pollution monitoring site where it can be erected at a height of 10 m while satisfying the “open grounds” criteria rather than compromising on anemometer height for the sake of closeness.

4. Conclusions

This study explores the capability of a simple modelling tool, SOSE, to identify the most representative wind site to be used for local-scale air pollution modelling applications when a choice of the wind site has to be made from a set of available wind sites. The results not only support important aspects stated in the WMO and EPA meteorological monitoring guidance but also provide quantitative model performance statistics for all available wind sites.

According to the results of this study, selecting wind data that is measured in close proximity is very important in reducing the prediction error of an air quality model, especially in situations where there is a great deal of spatial variability in the surface wind flows. However, colocation should not be a priority over choice of an open area, free from local obstacles such as trees and buildings, consistent with the WMO guidelines for wind data collection sites. If an open area cannot be found in close proximity for implementing the anemometer, national weather station data can be used with good accuracy. However, choosing the appropriate weather station should be done carefully. According to the results of this analysis, sites within 10 km gave reasonable model accuracy. However, not all the sites within 10 km radius performed well. The wind data that showed similar wind directions and a similar fraction of data from each direction to the most representative on site wind data (as can be seen from the wind roses) gave good predictions.

The intensity variations due to weakening of wind speeds do not make a significant difference to the prediction accuracy of the SOSE model as long as the consistency of the wind pattern is maintained (e.g., the LS and M sites). A wind site identified as a representative wind site by SOSE will provide representative wind data for any air quality model. If multiple wind sites are available, it is worth experimenting with wind data from all available sites to identify most representative wind data that characterise the dispersion of the pollutants to the site of interest. This may be particularly important in

areas of complex terrain as the winds are likely to be spatially inhomogeneous. The SOSE model is ideal for such an exercise as it is entirely dependent on representative wind data for good model performance. It is also computationally simple, is easy to implement, has no input requirements beyond wind speed and direction, and has been shown to be reliable for use in topographically complex areas. As such, the approach presented here can be expected to be transferrable to any urban area, irrespective of the topographical complexity or the nature of the road layout.

An analysis of wind data and air pollution model performance of this nature, using a simple tool such as SOSE, gives a clear basis for selecting representative wind data for air quality models. A valuable future exercise would be to verify these results using another statistical method such as artificial neural networks.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to acknowledge the New Zealand Transport Agency and the Foundation for Research, Science & Technology, New Zealand (“Healthy Urban Atmospheres” programme) for funding the Roadside data collection campaign. The authors would also acknowledge Sally Gray, Lou Reddish, and Nick Talbot from The National Institute of Water and Atmospheric Research (NIWA) Ltd., Auckland, New Zealand, for their technical support during the field campaign, Igor Drecki for doing the site map, and Stuart Grange for the MATLAB code for wind rose plots, the National Climate Database maintained by NIWA (<http://cliflo.niwa.co.nz/>) and “openair” open-source tools for air pollution data analysis (<http://www.openair-project.org/>).

References

- [1] C. Arciszewska and J. McClatchey, “The importance of meteorological data for modelling air pollution using ADMS-Urban,” *Meteorological Applications*, vol. 8, no. 3, pp. 345–350, 2001.
- [2] J. Wieringa, “Does representative wind information exist?” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 65, no. 1–3, pp. 1–12, 1996.
- [3] T. R. Oke, “Siting and exposure of meteorological instruments at urban sites,” in *Air Pollution Modeling and Its Application XVII*, C. Borrego and A.-L. Norman, Eds., pp. 615–631, Springer, New York, NY, USA, 2007.
- [4] T. R. Oke, “Initial guidance to obtain representative meteorological observations at urban sites. Instruments and methods of observation program,” IOM Report 81, WMO/TD 1250, World Meteorological Organization, Geneva, Switzerland, 2004, <https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-81/IOM-81-UrbanMetObs.pdf>.
- [5] World Meteorological Organization, *Guide to Meteorological Instrumentation and Methods of Observations*, Publications Board World Meteorological Organization (WMO), 2008.

- [6] U S Environmental Protection Agency, *Meteorological Monitoring Guidance for Regulatory Modeling Applications*, Office of Air Quality—Planning and Standards, 2000, <http://www.epa.gov/scram001/guidance/met/mmgrma.pdf>.
- [7] US Environmental Protection Agency, Federal Register, Rules and Regulations 70(216): 68243–68246, 2005, http://www.epa.gov/scram001/guidance/guide/appw_05.pdf.
- [8] W. F. Dabberdt, M. A. Carroll, D. Baumgardner et al., “Meteorological research needs for improved air quality forecasting,” *Bulletin of the American Meteorological Society*, vol. 85, no. 4, pp. 563–586, 2004.
- [9] K. N. Dirks, M. D. Johns, J. E. Hay, and A. P. Sturman, “A simple semi-empirical model for predicting missing carbon monoxide concentrations,” *Atmospheric Environment*, vol. 36, no. 39–40, pp. 5953–5959, 2002.
- [10] K. N. Dirks, M. D. Johns, J. E. Hay, and A. P. Sturman, “A semi-empirical model for predicting the effect of changes in traffic flow patterns on carbon monoxide concentrations,” *Atmospheric Environment*, vol. 37, no. 19, pp. 2719–2724, 2003.
- [11] P. Giambini, P. Salizzoni, L. Soulhac, and A. Corti, “Influence of meteorological input parameters on urban dispersion modelling for traffic scenario analysis,” in *Air Pollution Modeling and its Application XXI*, vol. 4, pp. 453–457, Springer, 2012.
- [12] L. Wang, N. Zhang, Z. Liu, Y. Sun, D. Ji, and Y. Wang, “The influence of climate factors, meteorological conditions, and boundary-layer structure on severe haze pollution in the Beijing-Tianjin-Hebei region during January 2013,” *Advances in Meteorology*, vol. 2014, Article ID 685971, 14 pages, 2014.
- [13] S. Gokhale and N. Raokhande, “Performance evaluation of air quality models for predicting PM₁₀ and PM_{2.5} concentrations at urban traffic intersection during winter period,” *Science of the Total Environment*, vol. 394, no. 1, pp. 9–24, 2008.
- [14] A. J. Manning, K. J. Nicholson, D. R. Middleton, and S. C. Rafferty, “Field study of wind and traffic to test a street canyon pollution model,” *Environmental Monitoring and Assessment*, vol. 60, no. 3, pp. 283–313, 2000.
- [15] J. L. Pearce, J. Beringer, N. Nicholls, R. J. Hyndman, and N. J. Tapper, “Quantifying the influence of local meteorology on air quality using generalized additive models,” *Atmospheric Environment*, vol. 45, no. 6, pp. 1328–1336, 2011.
- [16] E. Kovač-Andrić, T. Radanović, I. Topalović, B. Marković, and N. Sakač, “Temporal variations in concentrations of ozone, nitrogen dioxide, and carbon monoxide at Osijek, Croatia,” *Advances in Meteorology*, vol. 2013, Article ID 469786, 7 pages, 2013.
- [17] S. Righi, P. Lucialli, and E. Pollini, “Statistical and diagnostic evaluation of the ADMS-Urban model compared with an urban air quality monitoring network,” *Atmospheric Environment*, vol. 43, no. 25, pp. 3850–3857, 2009.
- [18] Auckland Air Emission Inventory, 2004, <http://www.arc.govt.nz/albany/fms/main/Documents/Environment/Pollution/airfacts6.pdf>.
- [19] I. Longley, S. Kingham, K. Dirks, E. Somervell, W. Pattinson, and A. Elangasinghe, “Detailed observations and validated modelling of the impact of traffic on the air quality of roadside communities,” NZ Transport Agency Research Report 516, 2013, <http://www.nzta.govt.nz/assets/resources/research/reports/516/docs/516.pdf>.
- [20] K. N. Dirks, A. Nanni, and V. I. Dirks, “Modelling and predicting urban atmospheric pollutants in the Aosta Valley region of Italy using a site-optimised model,” *Atmospheric Science Letters*, vol. 7, no. 1, pp. 15–20, 2006.
- [21] M. A. Elangasinghe, K. N. Dirks, N. Singhal, S. B. Costello, I. Longley, and J. A. Salmond, “A simple semi-empirical technique for apportioning the impact of roadways on air quality in an urban neighbourhood,” *Atmospheric Environment*, vol. 83, pp. 99–108, 2014.
- [22] D. P. Chock, “A simple line-source model for dispersion near roadways,” *Atmospheric Environment*, vol. 12, no. 4, pp. 823–829, 1978.
- [23] J. C. Chang and S. R. Hanna, “Air quality model performance evaluation,” *Meteorology and Atmospheric Physics*, vol. 87, no. 1–3, pp. 167–196, 2004.
- [24] S. R. Hanna, “Uncertainties in air quality model predictions,” *Boundary-Layer Meteorology*, vol. 62, no. 1–4, pp. 3–20, 1993.