Atmospheric Environment 144 (2016) 69–78

Contents lists available at ScienceDirect

Atmospheric Environment

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

Development of Land Use Regression models for particulate matter and associated components in a low air pollutant concentration airshed

Mila Dirgawati [a, *], Jane S. Heyworth [a], Amanda J. Wheeler [a, b], Kieran A. McCaul [c], David Blake [b], Jonathon Boeyen [b], Martin Cope [d], Bu Beng Yeap [e, f], Mark Nieuwenhuijsen [g], Bert Brunekreef [h], Andrea Hinwood [b]

[a] School of Population Health, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia
[b] School of Natural Sciences, Edith Cowan University, 270 Joondalup Drive, Joondalup, WA 6027, Australia
[c] WA Centre for Health & Ageing, The University of Western Australia, Level 6, Ainslie House, 48 Murray St, Perth, WA 6000, Australia
[d] CSIRO Marine and Atmospheric Research, PMB1 Aspendale, Vic 3195, Australia
[e] School of Medicine and Pharmacology, The University of Western Australia, Crawley, WA 6009, Australia
[f] Department of Endocrinology and Diabetes, Fremantle and Fiona Stanley Hospitals, Fremantle, WA 6959, Australia
[g] ISGlobal Center for Research in Environmental Epidemiology (CREAL), Barcelona Biomedical Research Park, 08003 Barcelona, Spain
[h] Institute for Risk Assessment Sciences, Utrecht University, P.O. Box 80178, 3508 TD Utrecht, The Netherlands

HIGHLIGHTS

- Development of LUR model in area with low air pollutant concentration is possible.
- Traffics are the main sources of airborne particulate matters and the elements.
- Most of LUR models explain more than 50% of the pollutant's spatial variability.
- Lack of specific predictor data for air pollutant limit the models' performance.

ARTICLE INFO

Article history:
Received 14 March 2016
Received in revised form
15 July 2016
Accepted 4 August 2016
Available online 5 August 2016

Keywords:
Land use regression (LUR) model
Air pollution
Particulate matter
PM elements

ABSTRACT

Perth, Western Australia represents an area where pollutant concentrations are considered low compared with international locations. Land Use Regression (LUR) models for PM_{10}, $PM_{2.5}$ and $PM_{2.5}$ Absorbance ($PM_{2.5}$Abs) along with their elemental components: Fe, K, Mn, V, S, Zn and Si were developed for the Perth Metropolitan area in order to estimate air pollutant concentrations across Perth. The most important predictor for PM_{10} was green spaces. Heavy vehicle traffic load was found to be the strongest predictor for $PM_{2.5}$Abs. Traffic variables were observed to be the important contributors for PM_{10} and $PM_{2.5}$ elements in Perth, except for $PM_{2.5}$ V which had distance to coast as the predominant predictor. Open green spaces explained more of the variability in the PM_{10} elements than for $PM_{2.5}$ elements, and population density was more important for $PM_{2.5}$ elements than for PM_{10} elements. The $PM_{2.5}$ and $PM_{2.5}$Abs LUR models explained 67% and 82% of the variance, respectively, but the PM_{10} model only explained 35% of the variance. The $PM_{2.5}$ models for Mn, V, and Zn explained between 70% and 90% of the variability in concentrations. PM_{10} V, Si, K, S and Fe models explained between 53% and 71% of the variability in respective concentrations. Testing the models using leave one-out cross validation, hold out validation and cross-hold out validation supported the validity of LUR models for PM_{10}, $PM_{2.5}$ and $PM_{2.5}$Abs and their corresponding elements in Metropolitan Perth despite the relatively low concentrations.

* Corresponding author.
E-mail address: mila.dirgawati@uwa.edu.au (M. Dirgawati).

http://dx.doi.org/10.1016/j.atmosenv.2016.08.013
1352-2310/

1. Introduction

Particulate matter (PM) is a heterogeneous mixture of suspended particles and varies in chemical composition and size (Liang, 2013). Long term exposures to PM mass with an aerodynamic diameter smaller than 10 μm (PM_{10}) and 2.5 μm ($PM_{2.5}$) respectively, have been associated with mortality, cardiovascular disease, lung cancer, and both chronic and acute respiratory diseases, including asthma, even at concentrations below ambient air quality standards (Andersen et al., 2012; Beelen et al., 2014; Cesaroni et al., 2014; Hoek et al., 2013; Raaschou-Nielsen et al., 2013).

The composition of PM may be important in determining the contribution of particular sources. A source apportionment study conducted in four Australian cities identified species-related sources of PM, including aluminium (Al) and silicon (Si) from crustal/soil dust; iron (Fe), copper (Cu), zinc (Zn) and Manganese (Mn) from motor vehicle emissions; potassium (K) from biomass burning emissions; and heavy metals (Fe, Cu, Zn) from industrial emissions (Chan et al., 2008; Larson et al., 2004).

There is some evidence that the composition of PM may also be important in determining health effects (Eeftens et al., 2014; Stanek et al., 2011). PM_{10}-associated zinc (Zn) has been associated with the risk of pneumonia among children in seven birth cohort studies in Europe (Fuertes et al., 2014). $PM_{2.5}$-associated nickel (Ni) and vanadium (V) have also been associated with daily mortality in time-series studies in two northern American cities (Zhou et al., 2010). Small decreases in lung function among young children have also been related with exposure to other PM-associated elements such as Cu, Fe, K, Si, sulfur (S) and Zn (Eeftens et al., 2014).

There are relatively little data available on the intra-urban speciation of PM_{10} and $PM_{2.5}$ for estimating exposures in large population-based health studies in Australia. Land Use Regression (LUR) models have been used to predict small-scale spatial variations in exposure to air pollutants including species of PM within cities in other locations (Gulliver and Briggs, 2011; Hoek et al., 2008; Zou et al., 2009). The recent European Study of Cohorts for Air Pollution Effects (ESCAPE) developed PM speciated LUR models for eight elements in 15 countries. Most of those LUR models explained a large fraction of the spatial variation within the study area with R^2 ranging from 50% to 79% (de Hoogh et al., 2013) and were used to investigate the associations with a range of health outcomes (Eeftens et al., 2014).

This paper describes the development of LUR models for PM_{10}, $PM_{2.5}$, $PM_{2.5Absorbance}$ ($PM_{2.5}$Abs), and PM-associated elements (Cu, Fe, K, V, S, Si, Zn, Ni, and Mn) for the Perth metropolitan area, Western Australia, an area with lower air pollutant concentrations compared with most European and North American cities. The models were then used to assign exposures to PM and PM elements for a cohort of older men in Perth, the Health in Men Study (HIMS) (Norman et al., 2009).

2. Methods

2.1. Area of study

The study was conducted in the Perth Metropolitan Area which is the capital of Western Australia. It is located on the Indian Ocean with the Darling Ranges to the East, and with two main waterways, the Swan and Canning rivers. Perth metropolitan area has an area of 6418 km^2 and the population was around 1.97 million in 2012. Its topography is mainly flat with an altitude of approximately 31.5 m above sea level and a Mediterranean climate (Yimin and Lyons, 2003).

2.2. Sampling sites selection

The monitoring site selection has been described elsewhere (Dirgawati et al., 2015). Twenty sites were selected to measure $PM_{2.5}$ and PM_{10}, comprising two regional background sites; eight urban background sites; and ten street sites based on the criteria described in the protocol for the ESCAPE study (http://www.escapeproject.eu/manuals/). One of the regional background sites was co-located at a monitoring station operated by the Western Australian Department of Environment Regulation for comparison with standard government monitoring. A reference site was also operated continuously throughout the sampling period to adjust for any temporal variability of particle concentrations between seasons. Fig. 1 shows a map of the sampling sites.

2.3. Sampling and analysis method

Following the ESCAPE protocol, sampling occurred over three seasons, summer, autumn and winter from January 31 to September 5, 2012, at 20 sites, shown in Fig. 1. Samples were collected from each site for a two-week period in each season, with a maximum of six sites being monitored concurrently. This resulted in four sampling periods per season. The reference site was operated continuously throughout 2012.

Harvard Impactors (MS&T, Air Diagnostics and Engineering Inc. Harrison, ME) were used to collect $PM_{2.5}$ and PM_{10} samples onto Teflon filters (37 mm 2 μm pore size PALL Life Sciences PTFE Membrane). Sampler flow rates were 10 L per minute ±5% and recorded before and after the collection. Samples were collected for 15 min every 2 h to prevent overloading so that a 42-hr sample was collected over two weeks. The sampling time and the mean flow rate values were then used to calculate the sample volume. All individual samples at all sites were checked to decide whether the samples were valid to precede subsequent analyses. Valid samples of PM mass are those with total sampling time of at least 67% of 42-hr over fourteen days were included in subsequent analyses. Those considered as invalid samples ($n = 9$) were removed from subsequent analyses.

Valid filters were weighed after being placed in the weighing room at a temperature of 23 ± 1 °C and relative humidity of 37 ± 2% for 48 h. The pre and post weighing was undertaken on a

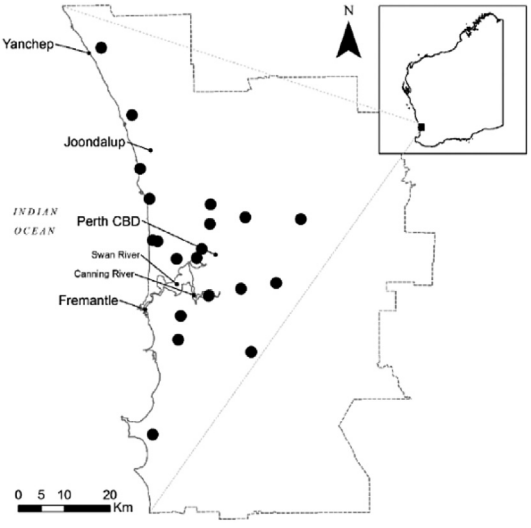

Fig. 1. Air sampling sites for PM_{10}, and $PM_{2.5}$ across Perth Metropolitan Area, Western Australia, 2012 (Dotted line indicates boundaries of the Perth Metropolitan Area).

microbalance to obtain mass values as described in the ESCAPE protocol (http://www.escapeproject.eu/manuals/). The mass was then divided by the sample volume to determine the corresponding concentrations of PM_{10} and $PM_{2.5}$. The PM_{coarse} concentrations were obtained by subtracting $PM_{2.5}$ from PM_{10} concentrations.

Reflectance was measured on all $PM_{2.5}$ post-weighed samples using a Smokestain reflectometer (EEL43M Smokestein Reflectometer (Diffusio System Ltd)). Reflectance was transformed into absorbance according to the International Standardization Organization method (1993). The measurements were limited to $PM_{2.5}$ filters as most of elemental carbon has been found in the $PM_{2.5}$ fraction (Eeftens et al., 2012). All PM mass and reflectance samples were prepared and analysed at the School of Natural Sciences laboratory, Edith Cowan University.

Filters were then analysed for 48 elements by energy dispersive X-ray fluorescence (ED–XRF). These analyses were undertaken by Cooper Environmental Services, Portland, Oregon, USA. Elemental concentrations were calculated by multiplying the reported mass per area of an element ($\mu g/cm^2$) with the exposed filter area (7.8 cm^2), subtracting the mean field blank and dividing by the individual filter's sample volume (de Hoogh et al., 2013).

The elements were selected for LUR model development based upon environmental emission characteristics, high percentage of detectability (more than 75% of detected samples), evidence of impact on human health (de Hoogh et al., 2013) and attribution of particular sources (Bukowiecki et al., 2010; de Hoogh et al., 2013; Zhang et al., 2015). Using these criteria, nine elements for Perth were identified, included Cu, Fe, and Zn for anthropogenic and traffic sources; K for biomass burning; S, and V for fossil fuel combustion; Si for soil and dust; Mn and Ni for industrial activities.

2.4. Quality assurance

Blanks and duplicate samples were collected for PM_{10} at the reference site only. Field blanks were collected at 19 time points throughout the annual sampling period. The duplicate samples were used to test the precision of the measurements and were assessed by calculating the absolute value of the difference between one instrument reading and the mean of the two, divided by that mean.

The limit of detection (LOD) was generated as three times the standard deviation of the field blanks. The LOD for PM mass was 0.56 $\mu g/m^3$. The LODs provided by the ED–XRF for V, Si, K, S, Fe, Mn, Zn, and Cu were 2.6×10^{-5} $\mu g/m^3$, 8.4×10^{-4} $\mu g/m^3$, 2.1×10^{-4} $\mu g/m^3$, 1.8×10^{-4} $\mu g/m^3$, 2.9×10^{-4} $\mu g/m^3$, 4.8×10^{-5} $\mu g/m^3$, and 7.5×10^{-5} $\mu g/m^3$, and 8.6×10^{-5} $\mu g/m^3$ respectively.

The measured concentrations of individual samples of PM_{10}, $PM_{2.5}$, PM_{coarse}, and $PM_{2.5}$Abs that were below the LOD were assigned a random value between 0 and the LOD. Elements were excluded from further evaluation if 25% or more of the samples were below the LOD. Concentration of individual samples below the LOD were not replaced with any other value, resulting in some negative values after the blank corrections (subtract the mean field blank from the sample values) were applied.

2.5. Calculation of annual average concentration

There were six Harvard Impactors available and hence the maximum number of sampling periods per season was four to enable the rotation of instruments to the varying locations. While the measurements were conducted to capture spatial variability, differences in the measured concentrations among the sites may exist because of temporal variations. Therefore, the concentrations at all sites across the three sampling seasons were adjusted using data from the reference site operated every two-week for the whole year 2012. A temporal correction factor for each two-week period within each season was calculated as the difference between the two weeks-specific-concentrations at the reference site and the annual average concentrations at the reference site. The correction factor was then subtracted from each measurement for the same two-week period to obtain a site specific adjusted concentration for each season. The adjusted seasonal concentrations were averaged to provide adjusted annual average concentrations for each site. These procedures were also applied to all elements. Annual averages for specific elements (Cu, Fe, K, V, S, Si, Zn, Ni, and Mn) were calculated when two or more seasons of data were available per site.

A scatterplot between unadjusted average concentrations and the adjusted concentrations including the R^2 and the linear regression equation were generated to evaluate the impact of temporal adjustment (data are not presented here). PM mass and the elements that were in poor agreement between the measured and corrected concentrations data were excluded from subsequent analyses.

Further, a correlation matrix of the measured concentrations of PM mass fractions and the elements was generated to investigate the relationships among these pollutants and differences in the source profiles.

2.6. Environmental predictor variables

Potential environmental predictors for developing the models were generated to cover a range of air pollutant sources that fit the local characteristics of Perth Metropolitan area and were categorised as: land use, population/household density and traffic variables.

The land use variables included buildings, industries, presence of water, proximity to water bodies, and open green spaces. The variables for buildings and industries were developed based on the Planning Land Use Classes (PLUCs) from the Valuer General's Office (VGO) of Western Australia Ministry for Planning for the year 2009, while data on water bodies were sourced from the 2009 Western Australia Land Information Authority (Landgate).

Industries were grouped into: (1) industrial facilities such as manufacturing/processing/fabrication, storage and distribution, and service industry; (2) commercial activities such as shops, retail outlets, offices/business, entertainment/recreational & cultural; (3) primary and rural activities such as farms & conservation areas; and (4) utilities such electricity, gas, water and waste services, transport, postal, and warehousing. For LUR model development, three definitions of industry variable were used: Industry-1 comprised: (1) industrial, (2) commercial, (3) primary/rural, and (4) utilities groups; Industry-2 comprised: (1) industrial, (3) primary/rural, and (4) utilities; and Industry-3 comprised: (1) industrial, and (2) commercial activities only.

The building variable represented all surrounding building types including residential buildings. The water body was characterized by sea, lakes and rivers. The amount of green spaces was represented by Normalized Difference Vegetation Index (NDVI), derived from the Landsat satellite data collected in 2012. The mesh block count from the Australian Bureau Statistics (ABS) for the year 2011 (www.abs.gov.au/websitedbs/censushoe.nsf) was used to derive the population and household density data. Both land use and population/household density variables were measured around the sampling sites within circular buffers with radii of 100 m, 300 m, 500 m, 1000 m and 5000 m to illustrate dispersion patterns of the pollutants being modelled and to capture the spatial variations of pollutant concentrations at both local and regional scales.

Traffic variables were measured in radii of 25 m, 50 m, 100 m,

300 m, 500 m and 1000 m to capture the local impact of potential traffic sources on air pollutant concentrations. Air pollutant concentrations decline exponentially with the distance to road, and decreases to the background levels behind a row of uninterrupted buildings (Batterman et al., 2010). Therefore, traffic variables without buffers such as traffic intensity on the nearest road and distance to major roads were also determined to account for the influence of nearby traffic emissions at the monitoring sites.

The traffic-related data such as number and type of vehicles on a given road, hierarchy of capacity of roads, length of roads and location of roads were obtained from the Main Roads WA (www.mainroads.wa.gov.au), collected for the year 2009. Main Roads WA monitors vehicle counts on selected roads using counters and calculates annual average daily traffic (AADT), the annual average number of vehicles travelling in both directions is adjusted for season and time of day. Main Road's WA also classifies the roads based upon its capacity including: primary distributor (>15,000 vehicles per day), distributor A (8001–15,000 vehicles per day), distributor B (6001–8000 vehicles per day), local distributor (3001–6000 vehicles per day), and access roads (<3000 vehicles per day). If no counters were present for a particular road class, the traffic counts from nearby roads of the same class in the road hierarchy were used to estimate the annual average traffic counts within each site's buffer. The nearest distance to a road was obtained by measuring the shortest distance from the monitoring site to the nearest road. Heavy duty vehicles included trucks and buses.

In total, 124 potential predictor variables were measured as summarized in the Supplementary Table S1. The generation of environmental predictor variables used in the development of LUR model have previously been described (Dirgawati et al., 2015). All GIS work was conducted using ArcGIS version 10.2 (ESRI Inc., 2013).

2.7. Model development

LUR models were developed to estimate the annual average concentrations of PM and the corresponding elements (dependent variable) using predictor variables at all monitoring sites. Prior to the models development, standardized predictor variables were generated by subtracting the mean for each predictor from each individual predictor data point and dividing by the SD. Thus, each standardized variable has a mean of zero and a SD of one. The coefficient generated in the final LUR model estimates the change in concentrations associated with a one SD change in the predictor.

Descriptive summaries and scatter plots between the adjusted annual average pollutant concentrations for each site and each predictor variable were used to develop an initial list of suitable predictors and monitoring sites for inclusion in the modelling.

Suitable predictors for model development were those: (1) with 75% values above zero; and (2) where the resultant slope was in the expected direction, as determined *a priori* (for example, green space is expected to reduce the particulate concentrations, while traffic activity is expected to elevate the concentrations). Predictors within smaller buffer sizes including industry and water within buffers of 100 m and 300 m and traffic within a buffer radii of 25 m were found to have a considerable number of zero values (>80%), and thus were excluded from the modelling. Among the 124 potential predictors, the total number of suitable predictors to be considered in the models ranged from 44 to 71 predictors. Limiting the number of predictors considered for LUR models reduce the risk of overfitting that occurs when large numbers of predictors are considered to explain concentrations at a relatively small number of monitoring sites (Wang et al., 2012, 2013).

The adjusted annual average concentration of PM_{10}, $PM_{2.5}$, and $PM_{2.5}$Abs at each site were checked to detect potential outliers. If the annual average concentration at a particular site was above the 95% percentile, the value was determined as an outlier. This site was then evaluated to decide whether it should be excluded from the modelling, based upon: (1) the stability of the model, i.e. if the parameter estimates of the model (adjusted R^2, direction effect of the predictors) with and without this site differed considerably; (2) the site location was not representing the site specific environmental exposure. As a result, one monitoring site was determined as an outlier, leaving nineteen sites to develop the LUR models.

The modelling procedure was based on manual stepwise selection techniques, following the procedures outlined by the ESCAPE protocol (http://www.escapeproject.eu/manuals/). Briefly, univariate models were run for all suitable predictor variables and the model with the highest adjusted R^2 and the expected slope direction for the predictor was used as the starting model for generation of the multivariate model. The remaining predictor variables were then added one at a time to the starting model. Variables were included if they complied with the following criteria: (1) increased the adjusted $R^2 \geq 1\%$, (2) the coefficient agreed with the predefined direction of effect, and (3) did not change the direction of effect for predictors already in the model.

During the LUR model development for each air pollutant, the univariate analysis might identify multiple starting models with similar adjusted R^2. This resulted in multiple appropriate LUR models for that specified pollutant. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for all models were then reviewed for selecting the final LUR model among the alternate models. The selected final LUR model was the model with smaller AIC and BIC values.

Further, to determine how much of the variability of each predictor contributed to the air pollutant concentrations, the R^2 of the nested model was subtracted from the R^2 of the final LUR model.

2.8. Model evaluation

The final models were reviewed for multicollinearity, influential observations, and autocorrelation using diagnostic statistics for multiple regression models. High multicollinearity for model predictors was determined based on Variance Inflation Factor (VIF) values of more than three. The influential observation was examined to ensure the model was not affected by one or more individual sites using Cook's D value above one. The cut-off and graphical plots between the observed and predicted values were also reviewed. Moran's I analysis was performed to investigate the spatial autocorrelation of the residuals of the final LUR models.

The performance of PM mass and the element models was evaluated using the leave-one-out cross validation. Given the relatively small number of monitoring sites and further application of eligible PM mass models to epidemiological studies in Perth, additional evaluation of PM mass models was conducted to investigate the true predictive ability of the models and stability of the predictors included in the models. The methods used, were hold-out validation and cross-holdout validation.

In the leave-one-out cross-validation method, evaluation models were developed using all but one of the measurement sites and the predicted concentrations were compared with the measured concentrations at the omitted site (Refaeilzadeh et al., 2009). The adjusted R^2 and the root mean square error (RMSE) between the predicted and observed concentrations were then calculated and compared with the original model and the corresponding standard deviation as measures of model performance (de Hoogh et al., 2013; Eeftens et al., 2012). Lower RMSE values typically indicate more stable models (Hoek et al., 2008; Mölter et al., 2010).

The hold-out validation used training and test sets in the evaluation procedure. The total number of suitable monitoring sites for

model development (nineteen sites) was divided into training and test sets. Ten sites were selected as training dataset for modelling, and the remaining nine sites were used for prediction outsides the training sites. The sites were randomly selected based on the strata of site types to ensure proportionate distribution of the street, urban background and regional background sites. The selection was repeated nine times to give 10 sets in total. The predictors variables included in the model for all sites were used to develop the training models. Refitting the same predictor in smaller subsets may have changed the direction of effect of predictors in the training model. The same criteria of predictors included in the model was used consistently to obtain the true predictive ability of the training models. Further, the squared Pearson-correlation coefficient, which is equivalent to R^2 between the measured and the predicted concentrations at the test sites was calculated and the stability of LUR model's structure was tested to measure the performance of these models at smaller subsets.

In the cross–hold out validation method, one individual site was successively left out, leaving eighteen sites for developing an evaluation model (Wang et al., 2016). The process was repeated nineteen times to obtain nineteen evaluation models. Each of the evaluation model was then used to predict the concentrations at the site that was not included in developing the model. This process was conducted for all evaluation models, resulting in the predicted concentrations across nineteen monitoring sites. The true hold out R^2 was calculated as the R^2 between the measured and the predicted concentrations at these nineteen external sites. The R^2 and the true hold out R^2 were then compared with the full sites LUR model. The stability of the model's structure was also tested by comparing the predictors included in the full sites model with those included in the training and evaluation models. All statistical analyses were undertaken using the statistical software STATA version 12.1 and 13.1 (StataCorp LP, Texas, USA).

3. Results

3.1. Descriptive statistics of measured air pollutant concentrations

Fifty-one valid samples (85%) were collected from nineteen sites during summer, spring and autumn, 2012. For the elements, the precision of the laboratory method had less than 10% variability demonstrating that the methods were reproducible. Some of elements ($PM_{2.5}$ S, $PM_{2.5}$ Cu, PM_{10} Ni and $PM_{2.5}$ Ni) were unsuitable for LUR modelling. The temporal corrections resulted in poor agreement between the measured and corrected data for S and Cu in $PM_{2.5}$. Both PM_{10} Ni and $PM_{2.5}$ Ni were also excluded from the modelling as more than 25% of the elements' samples had concentrations below the LOD.

Table 1 provides a summary of the descriptive statistics of the temporally adjusted annual concentrations for all PM and the selected elements for PM_{10} and $PM_{2.5}$ for the remaining nineteen monitoring sites. Summary statistics of the unadjusted measured concentration for the 48 elements of $PM_{2.5}$ and PM_{10} are presented in Supplementary Table S2 and Table S3, respectively.

The relationships between the annual averages of PM_{10}, PM_{coarse}, $PM_{2.5}$, and $PM_{2.5}$Abs and the selected elements are presented in Supplementary Table S4. A high correlation was observed for the association between PM_{10} and PM_{coarse} ($r = 0.95$), moderate correlation between $PM_{2.5}$ and $PM_{2.5}$ Abs ($r = 0.57$), and poor correlation between the PM_{10} and $PM_{2.5}$ ($r = -0.13$). For the PM_{10} elements, the highest correlation was observed between PM_{10} and S ($r > 0.5$), while poor correlations were observed between PM_{10} and the remaining elements. $PM_{2.5}$ was moderately correlated with K, V, Mn, Zn, and Fe with $0.45 < r < 0.53$. $PM_{2.5}$Abs was correlated with $PM_{2.5}$ Fe, Mn, and Zn; PM_{10} S and V.

3.2. LUR models for $PM_{2.5}$, PM_{10}, $PM_{2.5A}$ and the associated elements

LUR models were developed to predict the concentrations of $PM_{2.5}$, PM_{10} and their respective elements as well as $PM_{2.5}$Abs using 19 measurement sites across Perth. The LUR models for PM_{coarse}, Cu and PM_{10} Zn had poor predictive ability (adjusted $R^2 < 20\%$) and diagnostics, and are not presented. Fig. 2 illustrates the proportion of the spatial variability in the measured concentrations of $PM_{2.5}$, PM_{10}, and $PM_{2.5}$Abs and the selected elements explained by the LUR models.

The $PM_{2.5}$ and $PM_{2.5}$Abs models had acceptable predictive ability with an adjusted R^2 above 65%, while the PM_{10} model had much lower predictive ability. The LUR models captured a greater proportion of the spatial variability in the measured crustal elements (Si and K) and S in PM_{10} fraction than that in $PM_{2.5}$. The $PM_{2.5}$ elemental models explained more of variability in the spatial variability of industrial related source element (Mn) and vehicle source element (Zn). The variability of industrial/fuel oil combustion-element (V) in PM_{10} was better captured than that in the $PM_{2.5}$. A similar proportion was observed in the spatial variability of non-tailpipe vehicle source (Fe) in both PM_{10} and $PM_{2.5}$.

Complete description of the structure of the LUR models is presented in Supplementary Table S5, the summary of predictor variables captured by each LUR model and the corresponding unique contributions is illustrated in Table 2. For PM_{10}, the most important predictor was green spaces. In contrast, heavy vehicle traffic load was found to be the strongest predictor for $PM_{2.5}$Abs. Traffic variables were observed to be the important contributors for PM_{10} and $PM_{2.5}$ elements in Perth, except for $PM_{2.5}$ V which had distance to coast as the predominant predictor. Open green spaces explained more of the variability in the PM_{10} elements than for $PM_{2.5}$ elements, population density was more important for $PM_{2.5}$ elements than for PM_{10} elements.

The model diagnostics for $PM_{2.5}$ and $PM_{2.5}$Abs using the variance inflation factors (VIF), Cook's D and Moran's I were acceptable (Supplementary Table S6). The VIF and the Cook's D values suggest the models do not violate the collinearity and influential observation assumptions. The Moran's I for all PM mass fractions and reflectance ranged from 0.083 to 0.970 with a p-value > 0.05, representing no spatial autocorrelation of the residuals. The elemental models did not violate the general assumptions for the development of the LUR models except for PM_{10} Fe, PM_{10} Mn, and $PM_{2.5}$ V. Based upon evaluation during the model development, it was observed that the parameter estimates of the models for PM_{10} Fe, PM_{10} Mn, and $PM_{2.5}$ V with and without this influential site changed considerably. Therefore, the corresponding site was excluded from modelling for these elements. Among the LUR models for the PM_{10} and $PM_{2.5}$ elements, we found no spatial autocorrelation measured by Moran's I, except for $PM_{2.5}$ Zn.

Table 3 shows the results of the model evaluation procedures, including the leave-one-out cross validation, hold out validation and cross-hold out validation. The differences in the adjusted R^2 between the final models and the leave one out cross validation results for $PM_{2.5}$ was 17% and $PM_{2.5}$Abs was 15%, indicating the spatial predictive ability of both models are relatively good.

Plots of the measured and the predicted concentrations for both pollutants are shown in supplementary Figure S1, indicating an agreement between the measured and predicted concentrations.

The hold-out and cross-hold-out validation procedures results also indicated the stability of our $PM_{2.5}$ and $PM_{2.5}$Abs LUR models. The top four predictors in the final $PM_{2.5}$ model such as surface area of water body, open green spaces, population density, and traffic intensities were dominant in the $PM_{2.5}$ evaluation models (Supplementary Table S7). For $PM_{2.5}$Abs evaluation models, the

Table 1
Descriptive Summary of temporally adjusted annual average concentrations for PM_{10}, PMcoarse, $PM_{2.5}$, $PM_{2.5}$Abs, PM_{10} elements and $PM_{2.5}$ elements (N = 19).

Pollutant	Mean	Median	SD	Min	Max
PM mass ($\mu g/m^3$):					
PM_{10}	17.1	16.4	5.0	8.9	30.3
PM_{coarse}	12.4	10.8	5.5	4.2	26.4
$PM_{2.5}$	4.7	4.5	1.6	1.5	7.8
PM absorbance ($10^{-5}m^{-1}$): $PM_{2.5}$A	0.7	0.7	0.3	0.2	1.5
PM_{10} elements (ng/m^3):					
Si	3.9×10^{-1}	3.6×10^{-1}	2.4×10^{-1}	1.3×10^{-3}	7.9×10^{-1}
S	4.9×10^{-1}	4.5×10^{-1}	1.3×10^{-1}	3.4×10^{-1}	7.8×10^{-1}
K	1.8×10^{-1}	1.9×10^{-1}	4.8×10^{-2}	9.3×10^{-2}	2.8×10^{-1}
V	1.2×10^{-3}	1.1×10^{-3}	5.0×10^{-4}	6.0×10^{-4}	2.4×10^{-3}
Mn	5.5×10^{-3}	5.1×10^{-1}	3.6×10^{-3}	1.8×10^{-3}	1.7×10^{-2}
Fe	2.9×10^{-1}	2.5×10^{-1}	1.7×10^{-1}	6.4×10^{-2}	6.8×10^{-1}
Cu	6.2×10^{-3}	4.6×10^{-3}	4.4×10^{-3}	1.0×10^{-3}	1.7×10^{-2}
Zn	9.3×10^{-3}	7.5×10^{-3}	5.8×10^{-3}	7.0×10^{-4}	2.3×10^{-2}
$PM_{2.5}$ elements (ng/m^3):					
Si	8.8×10^{-2}	8.8×10^{-2}	5.1×10^{-2}	5.6×10^{-2}	2.4×10^{-1}
K	7.9×10^{-2}	7.9×10^{-2}	2.0×10^{-2}	5.0×10^{-2}	1.2×10^{-1}
V	7.0×10^{-4}	7.0×10^{-4}	3.0×10^{-4}	3.0×10^{-4}	1.2×10^{-3}
Mn	2.4×10^{-3}	2.4×10^{-3}	3.7×10^{-3}	6.0×10^{-4}	1.6×10^{-2}
Fe	8.1×10^{-2}	8.1×10^{-2}	4.6×10^{-2}	2.4×10^{-2}	1.8×10^{-1}
Zn	6.5×10^{-3}	6.5×10^{-3}	5.9×10^{-3}	1.8×10^{-3}	2.9×10^{-2}

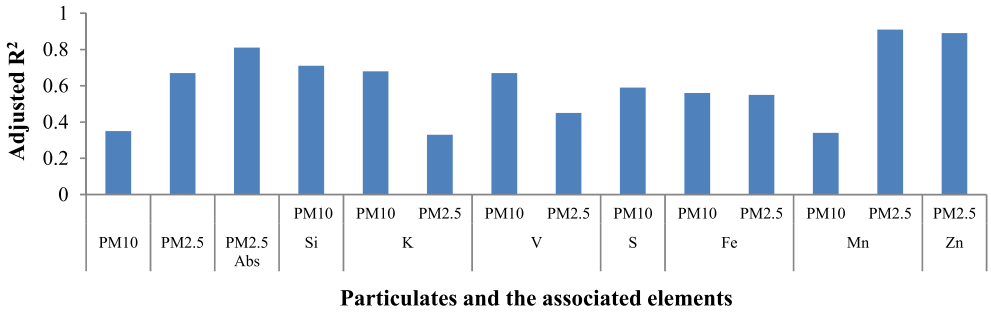

Fig. 2. Proportion of the spatial variability in the measured concentrations of $PM_{2.5}$, PM_{10}, and $PM_{2.5}$Abs and selected elements explained by the LUR models.

predictor matched. Traffic load, industry and commercial area, population densities, and proximity to coast were consistently captured by all evaluation models (Supplementary Table S8).

4. Discussion

4.1. Air pollutant concentrations

The annual mean concentrations of PM_{10} (17.1 $\mu g/m^3$), PM_{coarse} (12.4 $\mu g/m^3$), and $PM_{2.5}$ (4.7 $\mu g/m^3$) across Perth were below the National Environment Protection Measures (NEPM) for ambient air quality (20 $\mu g/m^3$ for PM_{10} and 8 $\mu g/m^3$ for $PM_{2.5}$) (NEPC, 1998) and the WHO annual mean air quality guidelines (20 $\mu g/m^3$ for PM_{10} and 10 $\mu g/m^3$ for $PM_{2.5}$) (WHO, 2005). The mean concentration of $PM_{2.5}$Abs was 0.67×10^{-5} m^{-1}. There is no corresponding guideline for this pollutant or for any of the PM elements. The high correlation between PM_{10} and PM_{coarse} reflects the fact that PM_{10} and PM_{coarse} are emitted from similar sources such as non-exhaust emissions and fugitive coarse dust (Keuken et al., 2013), while moderate correlation between $PM_{2.5}$ and $PM_{2.5}$ Abs, possibly reflecting that their concentrations are influenced by complex interaction between local meteorological conditions and the main source of these pollutants such as vehicles emissions and industrial activities (Keuken et al., 2013).

PM_{10} and PM_{coarse} annual average concentrations were at the lower end of those European cities participating in the ESCAPE study, which ranged from 14.8 to 43.1 $\mu g/m^3$ for PM_{10}; and 6.0–23.6 $\mu g/m^3$ for PM_{coarse}. Likewise, the $PM_{2.5}$ concentrations and the $PM_{2.5}$Abs concentrations were both lower than those of the ESCAPE cities, which were 8.3–29.3 $\mu g/m^3$ for $PM_{2.5}$; and 0.8–3.0 $\times 10^{-5}$ m^{-1} for $PM_{2.5}$Abs (Eeftens et al., 2012). The median concentrations of the PM elements were typically lower than the ESCAPE cities (de Hoogh et al., 2013) and the Calgary study (Zhang et al., 2015). Similarly to the PM mass and the reflectance measures, the concentrations of all PM mass elements in Perth were lower than those measured in European and North America cities (Ross et al., 2007). Possible reasons for these differences between the studies include: (1) air pollutant concentrations in Perth tend to disperse across the airshed due to its flat topography and strong ocean breezes (Yimin and Lyons, 2003); (2) the traffic intensities on major roads in Perth (<60,000 vehicles per day) are lower than the cities in the ESCAPE study, which are typically above 100,000 vehicles per day (EEA, 2011); and (3) road networks in the ESCAPE cities are denser than in Perth (EEA, 2011). The $PM_{2.5}/PM_{10}$ ratio is low (0.3), indicating that sea salt and wind blown dust are possibly responsible for the low $PM_{2.5}/PM_{10}$ ratio relative to the European cities (Eeftens et al., 2012).

4.2. Characterisation of surrounding environment and land use

LUR models identify surrounding environmental and land use characteristics that may help explain the variability in pollutant concentrations (Hoek et al., 2008). Traffic intensity of heavy vehicles in close proximity to the monitoring site was associated with

Table 2
Predictor variables and their unique contributions included in the LUR models for PM_{10}, $PM_{2.5}$, $PM_{2.5}$ Abs and their associated elements.

Predictor Variables	PM_{10}	$PM_{2.5}$	$PM_{2.5}$ Abs	PM_{10} elements					$PM_{2.5}$ elements					
				Si	K	V	S	Fe	Mn	Fe	Mn	V	Zn	K
Open green spaces	5000 m (33%)	1000 m (13%)		500 m (9%)	500 m (8%)	5000 m (3%)	5000 m (9%)	5000 m (9%)	1000 m (16%)		1000 m (2%)			
Population density		100 m (8%)	500 m (7%)							1000 m (15%)				
Housing density								100 m (5%)	100 m (8%)					
Building	100 m (4%)													
Industry			Industry-3[a] 5000 m (11%)								Industry-2[b] 1000 m (2%)	Industry-3[a] 1000 m (8%)		
Water body		5000 m (13%)	1000 m (1%)				5000 m (9.7%)	5000 m (13.8%)					5000 m (7%)	5000 m (10%)
Distance to water body		Coast (5%)	Coast (1%)			Coast (25%)						Coast (28%) River (1%)		River (5%)
Traffic intensity on nearest road		Dist A[c] (8%)			PDist (7%)					Dist A[c] (14%)				Dist A[c] (6%)
Heavy traffic intensity on nearest road	Any roads (27%)													
Traffic load			Any roads 500 m (25%)		Dist B[e] 500 m (6.6%)	Any roads 300 m (45%)	PDist 500 m (45%)	Dist B[e] 500 m (24%)	Dist B[e] 500 m (13%)	PDist[d] 1000 m (13%)	Any roads 1000 m (77%)		Any roads 1000 m (53%)	
Heavy vehicles traffic load		Dist B[e] 500 m (3%)					Dist B 1000 m (9%)							
Road length				Dist B[e] 500 m (15.2%)									PDist[d] 1000 m (3%)	Dist B[e] 300 m (20%)
Distance to road				PDist[d] (15%)	Dist B[e)] (1%) PDist[d] (2%)			PDist[d] (4%)						

[a] Industry-3: Industry variable – definition 3; comprised of industrial and commercial.
[b] Industry-2: Industry variable – definition 2; comprised of industrial, primary/rural and utilities.
[c] DistA: Distributor A road (8001–15,000 vehicles per day).
[d] PDist: Primary Distributor road (>15,000 vehicles per day).
[e] DistB: Distributor B road (6001–8000 vehicles per day).

Table 3
Summary of $PM_{2.5}$ and $PM_{2.5}$A models evaluation average R^2 Validation results of LUR models for $PM_{2.5}$ and $PM_{2.5}$Absorbance.

Air pollutant	Leave one out cross validation (n = 19)	Hold-out validation		Cross-hold-out validation	
		Training sets	Test sets	Training sets	Test sets
		(n = 10)	(n = 9)	(n = 18)	(n = 19)
$PM_{2.5}$	0.50	0.74	0.47	0.69	0.14
$PM_{2.5}$ Abs	0.67	0.74	0.73	0.82	0.61

increases in measured PM_{10} concentrations, contributing approximately 27% in the explained variability. More than 90% of the heavy vehicle fleets in Australia have diesel engines. Heavy vehicles account for around 25% of all road transport fuel consumed in Australia, thus they potentially contribute more to traffic exhaust (ABS, 2014). The LUR models for PM_{10} in European cities such as Manchester, London and Ruhr Area, have also included heavy vehicle traffic intensity in their models (Eeftens et al., 2012).

Traffic variables were included in all element models of PM_{10} and $PM_{2.5}$ fractions. Keuken et al. (2013) reported that Zn in particulates found in nearby roads originate from tire wear as its use as galvanised materials in rubber production and re-suspended road dust. Fe was found to be emitted from metal wear in the exhaust system walls as flakes of iron and it is considered as a good marker for brake wear emissions (Keuken et al., 2013). Fe can also be emitted from tailpipes, formed within the engine due to gas to particle conversion processes of the ferrocene, an agent to raise octane levels of diesel and gasoline fuel (Srimuruganandam et al., 2011). K is related to emissions from the ash fractions of diesel exhausts (Srimuruganandam et al., 2012), and Mn is emitted from brake lining dust (Grigoratos and Martini, 2015; Srimuruganandam et al., 2011). Traffic-related emissions are not specific sources for V and S in the PM_{10} fraction, as these elements are also emitted from fossil fuel combustion in industries (Murillo et al., 2013; Srimuruganandam et al., 2012). Most of the models' variables appear attributable to the resuspended road dust.

Asphalt roads comprise 95% mineral grains (Ca, Al, Si, Na, K) and 5% filler and binding materials (Srimuruganandam et al., 2012), suggesting Si may arise from road wear of asphalt roadways. Compared with the other elements' models that included many traffic intensity or traffic load predictors, the LUR model for Si uniquely included distance to major road and length of major road.

Therefore, the LUR model is likely to indicate the contribution of the re-suspended dust or paved road dusts to the concentrations of crustal elements in PM_{10}.

The models for $PM_{2.5}$, and $PM_{2.5}$Abs showed that population density within the smallest buffer sizes (100 m and 500 m) contributed to the increase in their concentrations. Population and housing density are associated with various anthropogenic sources including residential activities such the use of wood stoves and heaters, as well as tailpipe and non-tailpipe emissions of traffic servicing this area, as reported previously for other areas (Eeftens et al., 2012; Urman et al., 2014). Wood stoves and heaters are top emission sources of $PM_{2.5}$ in Australia as documented by the National Pollutant Inventory (NPI, 2014).

Population and housing density were also included in the Fe − $PM_{2.5}$ and Mn − PM_{10} LUR models, respectively. Fe is related to the dominant vehicle emissions, brake, and road wear, while Mn is mainly attributed to re-suspension of road dust reflecting the contribution of traffic servicing residential areas. The results are not surprising given higher number of street sites relative to the urban background and regional sites.

Our LUR models identify industrial activities (manufacturing, processing and fabrication) and commercial activities (shops, retails, offices, entertainment, recreational & cultural activities) located within 5000 m were modest predictors for $PM_{2.5}$Abs and $PM_{2.5}$ V, and minor predictors for $PM_{2.5}$ Mn. Such results support evidence that traffic variables are the major air pollution source for airborne $PM_{2.5}$ Abs and $PM_{2.5}$ in the Perth airshed. Refinery/residual oil combustion for industrial activities has been shown to be the indicator of V, and steel making has been the primary contributor of Mn (Chow and Watson, 2002). Zhang et al. (2015) also identified alternative predictors for Mn including industrial facilities. de Hoogh et al. (2013) identified sources of V to be industrial and fuel oil combustion-related. Both could also be the sources in Perth, given that oil combustion is one of the major air pollution sources in the industrial area in Perth (NPI, 2014).

The influence of the proportion of green space and water in an area resulted in the reduction of all of the PM and elemental concentrations within a range of buffer sizes. The LUR models for PM_{10} and all the associated elements characterised green spaces predominantly within buffer sizes of 500 m, 1000 m and 5000 m. Such associations between greenness and PM mass have also been found in other studies, suggesting plants and trees in open spaces play an important role in improving air quality and reducing the concentrations of particulates (Eeftens et al., 2012; McDonald et al., 2007). The observed association between the greenness and PM mass concentrations may be related to the proportion of green spaces that account for almost 15% of the total area of Perth and possibly lower traffic intensities in open green and water areas. Our PM models are consistent with LUR models for ESCAPE and for North American cities that included urban green and natural land use within large buffer sizes (1000 or 5000 m) (de Hoogh et al., 2013; Eeftens et al., 2012; Ross et al., 2007).

Off shore shipping are known to be indicators of V (Chow and Watson, 2002), thus shipping activity at ports may be one of the emission sources for V. Although the variable of ports was not considered in model development because there were 90% zero values around the monitoring sites, the V model included distance to coast, which contributed 25% and 28% to the increase in the concentrations of V in PM_{10} and $PM_{2.5}$ respectively. The LUR models, therefore, indicate the importance of the emission oil combustion of ship movement activities on the nearby coast to the V concentrations.

4.3. Spatial variability of air pollutant concentrations

We observed differences in the performance of each LUR model for explaining the spatial variability of particulates and their elements in Perth. The majority of the LUR models were able to explain more than 50% of the pollutant's spatial variability. There were some exceptions which included PM_{10}, PM_{10} Mn, $PM_{2.5}$ V, and $PM_{2.5}$ K which only explained between 30 and 45% of the spatial variability. The absence of specific predictor data for those species may limit the models' performance. For example, fugitive dust to explain the PM_{10} concentrations, refinery/residual oil industries and off shore shipping as the sources of V, steel industries as the primary contributor of Mn, and wood burning for K were not incorporated as potential environmental predictors in the modelling process. Consistent with previous studies in Europe and North America that suggest small variations in the measured element concentrations, lack of specific predictors and poor precision of measurements in areas with low concentrations are the main reasons for a poor R^2-model and R^2-LOOCV (de Hoogh et al., 2013; Eeftens et al., 2012; Ross et al., 2007).

The model R^2 of $PM_{2.5}$ was comparable with those in the ESCAPE study areas which ranged from 49% to 89% and were slightly higher than those reported in North America (Ross et al., 2007). The predictive power of the $PM_{2.5}$Abs model was also similar to the ESCAPE results (between 56% and 95%), while for the PM_{10} model, the predictive ability was lower than those in the ESCAPE (50%–90%) (Eeftens et al., 2012). The PM_{10} elemental models explained between 34% and 71% of the variability with only Fe, Zn and Cu performing poorly in comparison to the average R^2 for the ESCAPE models across all cities. The $PM_{2.5}$ elemental models explained 36%–90% of the concentration variability. Our Fe and K models performed less well in comparison with ESCAPE but our V and Zn models were able to explain a larger proportion of the variability than ESCAPE (de Hoogh et al., 2013).

The stability of our LUR models is also indicated by the results of the evaluation procedures. The differences in adjusted R^2 values between the model and the LOOCV were 17% for $PM_{2.5}$, 9% for PM_{10}, and 15% for $PM_{2.5}$Abs. For PM_{10}, $PM_{2.5}$ and $PM_{2.5}$Abs, the RMSE of LOOCV and models were smaller than the corresponding standard deviation. For both PM_{10} and $PM_{2.5}$ elements, the differences between the adjusted R^2-model and -LOOCV were within 15% suggesting the stability of the models, except for PM_{10} Mn (19%) and $PM_{2.5}$ K (26%). RMSE values that were obtained from LOOCV were found to be higher than those RMSE models. However, when the RMSE were assessed relative to the range of measured air pollutant concentrations, we found small differences between the LOOCV and the models i.e. 9% for $PM_{2.5}$ and 6% for $PM_{2.5}$Abs, demonstrating stable models.

From the hold out validation method, the $PM_{2.5}$ training models had larger average R^2 compared with the full-sites model, indicating over fitting. The results from the cross hold out validation method showed that the adjusted R^2 of the $PM_{2.5}$ models based on the full sites and 18 sites were similar. However, the true hold-out R^2 underestimated the models' predictive ability at the site that was not used for developing the $PM_{2.5}$ models. For $PM_{2.5}$Abs, the average R^2 of training models from the hold out validation and evaluation models from the cross hold out validation were similar to the adjusted R^2 of the full sites models. The training and evaluation models estimated the true predictive ability of the models at the external sites in the acceptable range, consistent with the LOOCV results which demonstrate the stability of the $PM_{2.5}$Abs model.

4.4. Limitations and generalizability

Our LUR models used a smaller number of sites compared with previous studies that had suggested a larger number of monitoring sites (>80 sites) were required (Basagaña et al., 2012). However, the

number of sites is consistent with the ESCAPE protocol. Overfitting may have occurred because of a large number of predictors (124 predictors) that were included in developing the models relative to the small number of sites (19 sites) (Basagaña et al., 2012; Wang et al., 2012). Despite the relatively small number of monitoring sites, this study has developed LUR models at an acceptable performance level with lower risk of overfitting as we implemented selection procedures for suitable predictors prior to modelling.

The availability of predictor data may also limit our model generalizability. Data on traffic intensities on relevant roads were obtained by using the available traffic counts. In addition, the concentrations of air pollutants were measured in 2012, while the predictor data such as household density were sourced from 2011 databases and industrial areas were obtained from 2009 databases. Data on predictors that are collected from the same year as when the monitoring occurred can optimize the predictive ability of models and enhance generalizability as the quality of predictor data at specific time periods can affect the results of LUR models (Hoek et al., 2008). However, the pattern of those predictors in Perth was relatively consistent across the five-year period, suggesting that the final LUR models are generally representative for capturing the spatial variability in air pollutants throughout the year 2012.

5. Conclusion

Despite the relatively low concentrations, LUR modelling for $PM_{2.5}$ and $PM_{2.5}$Abs and the elements of PM_{10} and $PM_{2.5}$ is possible for such locations. The LUR models characterised the local traffic related air pollution as the predominant source to explain the spatial variability of airborne particulate matter and the associated elements in Metropolitan Perth. This study represents one of the few LUR studies investigating PM elements and it will assist researchers in assessing the health impacts of the components of PM as well as PM size fractions.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.atmosenv.2016.08.013.

References

ABS, Australian Bureau of Statistics, 2014. In: Motor Vehicle Census, Australia, Retrieved 15 March, 2015. http://www.abs.gov.au/ausstats/abs@.nsf/mf/9309.0.

Andersen, Z.J., Bønnelykke, K., Hvidberg, M., Jensen, S.S., Ketzel, M., Loft, S., Sørensen, M., Tjønneland, A., Overvad, K., Raaschou-Nielsen, O., 2012. Long-term exposure to air pollution and asthma hospitalisations in older adults: a cohort study. Thorax 67 (1), 6–11.

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., Künzli, N., 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos. Environ. 54, 634–642. http://dx.doi.org/10.1016/j.atmosenv.2012.01.064.

Batterman, S.A., Zhang, K., Kononowech, R., 2010. Prediction and analysis of near-road concentrations using a reduced-form emission/dispersion model. Environ. Health 9, 29–39. http://dx.doi.org/10.1186/1476-069x-9-29.

Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z.J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet 383 (9919), 785–795.

Bukowiecki, N., Lienemann, P., Hill, M., Furger, M., Richard, A., Amato, F., Prévôt, A., Baltensperger, U., Buchmann, B., Gehrig, R., 2010. PM10 emission factors for non-exhaust particles generated by road traffic in an urban street canyon and along a freeway in Switzerland. Atmos. Environ. 44 (19), 2330–2340.

Cesaroni, G., Forastiere, F., Stafoggia, M., Andersen, Z.J., Badaloni, C., Beelen, R., Caracciolo, B., de Faire, U., Erbel, R., Eriksen, K.T., Fratiglioni, L., Galassi, C., Hampel, R., Heier, M., Hennig, F., et al., 2014. In: Long Term Exposure to Ambient Air Pollution and Incidence of Acute Coronary Events: Prospective Cohort Study and Meta-analysis in 11 European Cohorts from the ESCAPE Project, vol. 348.

Chan, Y.-C., Cohen, D.D., Hawas, O., Stelcer, E., Simpson, R., Denison, L., Wong, N., Hodge, M., Comino, E., Carswell, S., 2008. Apportionment of sources of fine and coarse particles in four major Australian cities by positive matrix factorisation. Atmos. Environ. 42 (2), 374–389.

Chow, J.C., Watson, J.G., 2002. Review of PM2. 5 and PM10 apportionment for fossil fuel combustion and other sources by the chemical mass balance receptor model. Energy & Fuels 16 (2), 222–260.

de Hoogh, K., Wang, M., Adam, M., Badaloni, C., Beelen, R., Birk, M., Cesaroni, G., Cirach, M., Declercq, C., Dedele, A., 2013. Development of land use regression models for particle composition in twenty study areas in Europe. Environ. Sci. Technol. 47 (11), 5778–5786.

Dirgawati, M., Barnes, R., Wheeler, A.J., Arnold, A.-L., McCaul, K.A., Stuart, A.L., Blake, D., Hinwood, A., Yeap, B.B., Heyworth, J.S., 2015. Development of land use regression models for predicting exposure to no 2 and no x in metropolitan Perth, western Australia. Environ. Model. Softw. 74, 258–267.

EEA, European Environment Agency, 2011. In: Size of the Vehicle Fleet (TERM 032) - Assessment Retrieved 28 April, 2015, from. http://www.eea.europa.eu/data-and-maps/indicators/size-of-the-vehicle-fleet/size-of-the-vehicle-fleet-2.

Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dėdelė, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., et al., 2012a. Development of land use regression models for PM2.5, PM2.5 absorbance, PM10 and PMcoarse in 20 european study areas; results of the ESCAPE project. Environ. Sci. Technol. 46 (20), 11195–11205. http://dx.doi.org/10.1021/es301948k.

Eeftens, M., Hoek, G., Gruzieva, O., Mölter, A., Agius, R., Beelen, R., Brunekreef, B., Custovic, A., Cyrys, J., Fuertes, E., 2014. Elemental composition of particulate matter and the association with lung function. Epidemiology 25 (5), 648–657.

Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dėdelė, A., Eriksen, K., et al., 2012b. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 – results of the ESCAPE project. Atmos. Environ. 62, 303–317. http://dx.doi.org/10.1016/j.atmosenv.2012.08.038.

Fuertes, E., MacIntyre, E., Agius, R., Beelen, R., Brunekreef, B., Bucci, S., Cesaroni, G., Cirach, M., Cyrys, J., Forastiere, F., 2014. Associations between particulate matter elements and early-life pneumonia in seven birth cohorts: results from the ESCAPE and TRANSPHORM projects. Int. J. Hyg. Environ. Health 217 (8), 819–829.

Grigoratos, T., Martini, G., 2015. Brake wear particle emissions: a review. Environ. Sci. Pollut. Res. Int. 22, 2491–2504. http://dx.doi.org/10.1007/s11356-014-3696-8.

Gulliver, J., Briggs, D., 2011. STEMS-Air: a simple GIS-based air pollution dispersion model for city-wide exposure assessment. Sci. Total Environ. 409 (12), 2419–2429. http://dx.doi.org/10.1016/j.scitotenv.2011.03.004.

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. Atmos. Environ. 42 (33), 7561–7578.

Hoek, G., Krishnan, R.M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., Kaufman, J.D., 2013. Long-term air pollution exposure and cardio- respiratory mortality: a review. [journal article]. Environ. Health 12 (1), 1–16. http://dx.doi.org/10.1186/1476-069x-12-43.

Keuken, M.P., Moerman, M., Voogt, M., Blom, M., Weijers, E.P., Röckmann, T., Dusek, U., 2013. Source contributions to PM2.5 and PM10 at an urban background and a street location. Atmos. Environ. 71, 26–35. http://dx.doi.org/10.1016/j.atmosenv.2013.01.032.

Larson, T., Gould, T., Simpson, C., Liu, L.-J.S., Claiborn, C., Lewtas, J., 2004. Source apportionment of indoor, outdoor, and personal PM2. 5 in Seattle, Washington, using positive matrix factorization. J. Air & Waste Manag. Assoc. 54 (9), 1175–1187.

Liang, J., 2013. Chemical Modeling for Air Resources Fundamentals, Applications, and Corroborative Analysis *Chapter 9-Particulate Matter*. Elsevier Science, Burlington, pp. 189–219.

McDonald, A.G., Bealey, W.J., Fowler, D., Dragosits, U., Skiba, U., Smith, R.I., Donovan, R.G., Brett, H.E., Hewitt, C.N., Nemitz, E., 2007. Quantifying the effect of urban tree planting on concentrations and depositions of PM10 in two UK conurbations. Atmos. Environ. 41 (38), 8455–8467. http://dx.doi.org/10.1016/j.atmosenv.2007.07.025.

Mölter, A., Lindley, S., de Vocht, F., Simpson, A., Agius, R., 2010. Modelling air pollution for epidemiologic research — Part I: a novel approach combining land use regression and air dispersion. Sci. Total Environ. 408 (23), 5862–5869. http://dx.doi.org/10.1016/j.scitotenv.2010.08.027.

Murillo, J.H., Roman, S.R., Marin, Rojas, F., J., Ramos, A.C., Jimenez, S.B., Gonzalez, B.C., Baumgardner, D.G., 2013. Chemical characterization and source apportionment of PM10 and PM2.5 in the metropolitan area of Costa Rica, Central America. Atmos. Pollut. Res. 4 (2), 181–190. http://dx.doi.org/10.5094/APR.2013.018.

NEPC, National Environment Protection Council, 1998. In: National Environment Protection (Ambient Air Quality) Measure Retrieved 15 December 2015. http://www.scew.gov.au/nepms/ambient-air-quality.

Norman, P.E., Flicker, L., Almeida, O.P., Hankey, G.J., Hyde, Z., Jamrozik, K., 2009. Cohort profile: the health in men study (HIMS). Int. J. Epidemiol. 38 (1), 48–52.

NPI, National Pollutant Inventory, 2014. In: 2013/2014 Data within Western Australia - All Substances from All Sources Retrieved 3 August 2015. http://www.npi.gov.au/npidata/.

Raaschou-Nielsen, O., Andersen, Z.J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M.J., Brunekreef, B., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., et al.,

2013. Air pollution and lung cancer incidence in 17 european cohorts: prospective analyses from the european study of cohorts for air pollution effects (ESCAPE). Lancet Oncol. 14 (9), 813–822. http://dx.doi.org/10.1016/S1470-2045(13)70279-1.

Refaeilzadeh, P., Tang, L., Liu, H., 2009. Cross-validation Encyclopedia of Database Systems. Springer, pp. 532–538.

Ross, Z., Jerrett, M., Ito, K., Tempalski, B., Thurston, G.D., 2007. A land use regression for predicting fine particulate matter concentrations in the New York City region. Atmos. Environ. 41 (11), 2255–2269.

Srimuruganandam, B., Nagendra, Shiva, M, S., 2011. Chemical characterization of PM10 and PM2.5 mass concentrations emitted by heterogeneous traffic. Sci. Total Environ. 409 (17), 3144–3157. http://dx.doi.org/10.1016/j.scitotenv.2011.04.042.

Srimuruganandam, B., Nagendra, Shiva, M, S., 2012. Source characterization of PM10 and PM2.5 mass using a chemical mass balance model at urban roadside. Sci. Total Environ. 433, 8–19. http://dx.doi.org/10.1016/j.scitotenv.2012.05.082.

Stanek, L.W., Sacks, J.D., Dutton, S.J., Dubois, J.-J.B., 2011. Attributing health effects to apportioned components and sources of particulate matter: an evaluation of collective results. Atmos. Environ. 45 (32), 5655–5663.

Urman, R., Gauderman, J., Fruin, S., Lurmann, F., Liu, F., Hosseini, R., Franklin, M., Avol, E., Penfold, B., Gilliland, F., 2014. Determinants of the spatial distributions of elemental carbon and particulate matter in eight Southern Californian communities. Atmos. Environ. 86, 84–92.

Wang, M., Beelen, R., Basagana, X., Becker, T., Cesaroni, G., de Hoogh, K., Dedele, A., Declercq, C., Dimakopoulou, K., Eeftens, M., Forastiere, F., Galassi, C., Gražulevičienė, R., Hoffmann, B., Heinrich, J., et al., 2013. Evaluation of land use regression models for NO2 and particulate matter in 20 european study areas: the ESCAPE project. Environ. Sci. Technol. 47 (9), 4357–4364. http://dx.doi.org/10.1021/es305129t.

Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G., Brunekreef, B., 2012. Systematic evaluation of land use regression models for NO2. Environ. Sci. Technol. 46 (8), 4481–4489. http://dx.doi.org/10.1021/es204183v.

Wang, M., Brunekreef, B., Gehring, U., Szpiro, A., Hoek, G., Beelen, R., 2016. A new technique for evaluating land-use regression models and their impact on health effect estimates. Epidemiology 27 (1), 51–56.

WHO, W.H.O, 2005. In: WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide.

Yimin, M., Lyons, T.J., 2003. Recirculation of coastal urban air pollution under a synoptic scale thermal trough in Perth, Western Australia. Atmos. Environ. 37 (4), 443–454. http://dx.doi.org/10.1016/S1352-2310(02)00926-3.

Zhang, J.J., Sun, L., Barrett, O., Bertazzon, S., Underwood, F.E., Johnson, M., 2015. Development of land-use regression models for metals associated with airborne particulate matter in a North American city. Atmos. Environ. 106, 165–177.

Zhou, J., Ito, K., Lall, R., Lippmann, M., Thurston, G., 2010. Time-series analysis of mortality effects of fine particulate matter components in Detroit and Seattle. Environ. health Perspect. 119 (4), 461–466.

Zou, B., Wilson, J.G., Zhan, F.B., Zeng, Y., 2009. Air pollution exposure assessment methods utilized in epidemiological studies. J. Environ. Monit. 11 (3), 475–490.