Gunawardena, Janaka, Ziyath, Abdul Mohamed, Egodawatta, Prasanna, Ayoko, Godwin A., & Goonetilleke, Ashantha
(2014)
Mathematical relationships for metal build-up on urban road surfaces based on traffic and land use characteristics.
Chemosphere, *99*, pp. 267-271.

Mathematical relationships for metal build-up on urban road surfaces based on traffic and land use characteristics

Janaka Gunawardena[a,b], Abdul M. Ziyath[b,c], Prasanna Egodawatta[b], Godwin A. Ayoko[b], Ashantha Goonetilleke[b*]

[a]Toowoomba Regional Council, PO Box 3021, Toowoomba Village Fair, Queensland 4350, Australia

[b]Science and Engineering Faculty, Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia.

[c]Department of Chemical and Process Engineering, Faculty of Engineering, University of Peradeniya, Sri Lanka.

E-mail: mjajanaka@gmail.com; mohamed.ziyath@qut.edu.au; p.egodawatta@qut.edu.au; g.ayoko@qut.edu.au; a.goonetilleke@qut.edu.au

*Corresponding Author: Ashantha Goonetilleke; Tel: 61 7 3138 1539; Fax: 61 7 3138 1170
Email: a.goonetilleke@qut.edu.au

Abstract
The study investigated the influence of traffic and land use parameters on metal build-up on urban road surfaces. Mathematical relationships were developed to predict metals originating from fuel combustion and vehicle wear. The analysis undertaken found that nickel and chromium originate from exhaust emissions, lead, copper and zinc from vehicle wear, cadmium from both exhaust and wear and manganese from geogenic sources. Land use does not demonstrate a clear pattern in relation to the metal build-up process, though its inherent characteristics such as traffic activities exert influence. The equation derived for fuel related metal load has high cross-validated coefficient of determination (Q^2) and low Standard Error of Cross-Validation (SECV) values indicates that the model is reliable, while the equation derived for wear-related metal load has low Q^2 and high SECV values suggesting its use only in preliminary investigations. Relative Prediction Error values for both equations are considered to be well within the error limits for a complex system such as an urban road surface. These equations will be beneficial for developing reliable stormwater treatment strategies in urban areas which specifically focus on mitigation of metal pollution.

Keywords: Metal build-up; Stormwater pollutant processes; Stormwater quality; Traffic emissions; Traffic pollutants; Urban road surfaces

1. Introduction
Stormwater transports a range of potentially toxic metal ions deposited on urban impervious surfaces, particularly from road surfaces, to receiving waters causing adverse aquatic ecosystem health impacts (Herngren et al., 2006). Effective management of stormwater related metal pollution requires accurate estimation of metal loads present on road surfaces based on an in-depth understanding of the metal build-up process. Though solids build-up process has been widely understood, the transferability of the knowledge to specific pollutants such as metals is very limited (Liu et al., 2012).

Egodawatta et al. (2013) developed a mathematical model to replicate the metal build-up process based on antecedent dry days. However, the model did not consider the specific influence exerted by widely acknowledged major anthropogenic sources of metals on road surfaces, namely traffic and land use activities (Goonetilleke et al., 2009; Mahbub et al., 2010). This constrains its use in the development of effective traffic and land use related pollution mitigation strategies. Accordingly, the aims of this study were to: (1) investigate the influence of traffic and land use characteristics in the metal build-up process on urban road surfaces; and (2) develop quantitative mathematical relationships to predict metal loads in the build-up on road surfaces based on traffic and land use parameters. The outcomes from this study will contribute to the evaluation of metal pollution in the urban environment from future changes in traffic and land use characteristics and for the development of reliable stormwater treatment strategies in urban areas.

2. Materials and methods

Eleven road sites were selected in the Gold Coast region, Queensland, Australia with the study sites encompassing variations in traffic and land use characteristics (Fig. S1 in Supplementary Information). The build-up samples were collected from 2.0 m x 1.5 m plot areas in the middle of the traffic lane using the wet and dry vacuuming system described by Mahbub et al. (2011), which consisted of a domestic vacuum cleaner fitted with a water filtration system. The selected plots were firstly dry vacuumed to collect most of the dust samples, and then wet vacuumed after spraying deionized water on the plots at 2 bar pressure for three minutes in order to collect the remaining fine particles. Prior to the sample collection, the procedure was tested under laboratory conditions and was found to be 97.4% efficient (Refer to Supplementary Information).

The samples collected were transported and stored in the laboratory under prescribed conditions until the analysis of the following metals commonly present on urban road surfaces was undertaken: cadmium (Cd), chromium (Cr), copper (Cu), lead (Pb), manganese (Mn), nickel (Ni) and zinc (Zn), using Method 200.8 for Inductively Coupled Plasma-Mass Spectroscopy (ICP-MS) (US-EPA, 1994) with TraceSELECT (Product No. 54704) as the certified reference material. The detection limits for ICP-MS for the selected metals were in the range of 0.001 to 0.005mg/L. Additionally, the total solids (TS) was determined using Methods 2540C and 2540D (APHA 2004).

Pollutant build-up on road surfaces is influenced by traffic factors such as traffic volume, congestion and vehicle mix (EPASGV, 1999), and land use characteristics within the vicinity of the site (Goonetilleke et al., 2009). In this context, the following surrogate indicators were selected to represent these influential factors in the analysis: annual average daily traffic volume (ADT_to) as a surrogate for traffic volume; volume to capacity ratio (V/C) as a surrogate for congestion and total heavy duty traffic volume (ADT_hv) as a surrogate for vehicle mix. Relevant data was obtained by undertaking classified traffic counts at the respective study sites. Surrogates that represented land use factors were the percentage of, industrial (I%), commercial (C%) and residential (R%) land use within 1 km radius from the sampling points. The data analysis was conducted using multivariate methods including principal component analysis (PCA), factor analysis (FA) and multiple linear regression analysis (MLR).

PCA transforms a large set of variables into an orthogonal set of principal components. PCA transforms original variables to orthogonal principal components (PCs) so that highest variance is associate to first few PCs. This results in the reduction in the number of variables, thereby facilitating effective interpretation of the data set. PCA outcomes are often presented as biplots, which enable the identification of the underlying relationships between objects and variables (Mostert et al., 2010). A detailed description of PCA can be found elsewhere (Adams, 1995). Similarly, FA uses few factors to explain the correlation between the variables By observing the characteristics of the variables correlated to a factor, it is possible to explain what a particular factor represents (Abdi, 2003). MLR is a regression technique that is often used to develop mathematical relationships for a dependent variable based on a number of independent variables (Ni et al., 2001). In this study, FA was performed using StatistiXL software (v. 1.8, 2008, statistiXL,Broadway–Nedlands, Australia) while MATLAB R2009b (Mathworks Inc, Natick, MA, USA) was used for MLR and PCA analysis.

3. Results and Discussions

3.1 Factor Analysis (FA)

FA was performed on the raw data matrix (Table 1) using principal component extraction method with orthogonal VARIMAX rotation technique, which results in factors being strongly correlated to a specific set of variables, while weakly correlated with other variables as a result of rotating the original factors (Egodawatta et al. 2013). This simplifies the interpretation of a complex data set as each variable is primarily associated with a specific factor (Abdi, 2003). The factors were extracted based on the initial eigenvalue criteria ≥ 1 and the results are presented in Table 2. It was hypothesised that metals with the same source of origin and build-up process are grouped under the same factor.

As evident in Table 2, Cr and Ni are associated with Factor 1 because they have relatively higher loadings in Factor 1 compared to the other factors. Similarly, Pb, Zn and Cu have higher loadings in Factor 2, while Mn has higher loading on Factor 3. This suggests that the source and build-up process for Pb, Cu and Zn are different to those for Cr and Ni, while the source and build-up process for Mn is different to these two metal groups. Furthermore, Cd has a loading of 0.29 on Factor 1, 0.11 on Factor 2 and a relatively high negative loading on Factor 3. This suggests that the source and build-up process for Cd would be different from those of the other metals.

3.2 Exploratory Principal Component Analysis (PCA)

To facilitate visual display and interpretation of the results, PCA were separately performed on standardised data matrices consisting of: (1) metal loads and traffic variables, and (2) metal loads and land use variables (Table 1). As shown in Fig. 1(a), when the metal load vectors are projected against PC1, all metals except Mn are strongly associated with PC1 suggesting that the primary source of Mn is different from that of the other metals. Mn is likely to be contributed to the build-up primarily by geogenic sources since the V/C vector, which is a highly influential traffic variable on PC1 is not correlated with Mn on PC1. In contrast, the rest of the metals have positive correlations with V/C when the corresponding vectors are projected on PC1, suggesting that they primarily originate from traffic sources.

The ‘traffic-related metals’ are discriminated on PC2 into two groups with Cu, Pb and Zn projected against the negative PC2 axis and Cr and Ni projected against the positive PC2 axis. Although Cd has a low negative PC2 loading, it does not clearly belong to either group. Similar conclusions were derived from FA, thus strengthening the PCA outcomes.

These observations clearly suggest that the two groups of metals originate from different primary traffic sources, while Cd may be contributed by both of these sources. It should also be noted that these metals could potentially originate from other sources such as industrial activities. However, the Gold Coast region predominantly consists of commercial and residential land uses with only light industries that have neither high fuel usage nor other significant metal generating activities. As such, it can be hypothesised that vehicular traffic is the major source of metals present on the road surfaces at the study sites.

Past studies have identified vehicle exhaust and vehicle component wear such as wear of brakes, tyres and engine components as the two major sources of traffic generated metals (e.g. Sansalone et al., 1997; Ball et al., 1998). Pb was used as the reference to discriminate between the metals contributed by exhaust and those by wear since the traffic source of Pb is likely to be vehicle component wear because leaded fuel has been phased out in Australia more than a decade ago. Therefore, it can be concluded that Pb, Cu and Zn are mainly contributed to road surfaces by the wear of vehicle brakes and tyres whereas Cr and Ni are primarily contributed by vehicle exhaust. Cd is contributed by both wear and exhaust. These conclusions are in agreement with the composition of vehicle exhaust and vehicle wear as reported in past research studies such as Galvagno et al.,(2002) and Mitrović et al., (2012).

This conclusion was further strengthened by the correlation of V/C vector with Pb, Cu and Zn and the correlation of both V/C and ADT_to with Cr and Ni in the PCA biplot (Fig. 1a). Frequent stops associated with traffic congestion can result in increased brake and tyre wear. In contrast, the smooth flow of traffic would result in only limited wear of vehicle components compared to frequent braking. Hence, the V/C vector is correlated with Pb, Cu and Zn, while ADT_to is not. On the other hand, vehicle exhaust increases with increased traffic volume and congestion, hence the correlation of V/C and ADT_to with Cr and Ni.

According to Fig. 1(b), both Mn and R% have negative loadings on PC1, while the other metals along with I% and C% have positive loadings. This further suggests that Mn originates primarily from geogenic sources, which are the main sources of metals in residential areas (Singh and Gilkes, 1992).

According to the relationships evident on the PC biplot, Mn is perpendicular to R% indicating that Mn does not have any correlation with R%. Similarly, I% correlates only with the Ni vector in contrast to the strong correlation of C% with Zn, Pb, Cu, Cd and Cr on PC1. In general, metals originating from traffic sources are strongly correlated with C% and weakly correlated with I%. This can be attributed to the fact that commercial areas have higher traffic activities compared to industrial and residential areas. Therefore, it can be concluded that it is not strictly land use that influences metal build-up, but rather it is the traffic characteristics. However, it is hypothesised that land use implicitly influences traffic characteristics.

In summary, the outcomes of PCA confirmed that metals in build-up are discriminated on the basis of their sources. Egodawatta et al. (2013) found that the build-up process of metals originating from the same source is similar, but different to that for metals originating from a different source. Hence, it is hypothesised that the build-up process for Pb, Cu and Zn are similar, but different to that for Cr and Ni.

3.3 Mathematical relationships for metal build-up based on traffic and land use characteristics

Based on the outcomes of PCA and FA discussed above, mathematical relationships between metals in build-up and influential traffic and land use factors were developed considering their sources of origin. Accordingly, Cr and Ni were grouped as fuel related metals (FM) and Pb, Zn and Cu as wear related metals (WM). Cd and Mn were not included in this analysis since Cd was present in relatively low quantity compared to other metals and Mn was deemed as originating from geogenic sources and outside the scope of this study.

Mathematical relationships for FM and WM were developed for their cumulative loads using Multiple Linear Regression Analysis (MLR). Prior to MLR, FM and WM were normalised with respect to the total solids load (TS), as research literature has identified the strong relationship between TS and metal loads (Helmreich et al. 2010). Accordingly, the dependent variables were (FM/TS) and (WM/TS), while the predictor variables were land use and traffic variables as shown below:

$$\left(\frac{\mathrm{FM}}{\mathrm{TS}}\right) = \sum\left(\frac{\mathrm{Cr\ load} + \mathrm{Ni\ load}}{\mathrm{TS\ load}}\right) = f\left(\mathrm{ADT\_to}, \mathrm{ADT\_hv}, \frac{\mathrm{V}}{\mathrm{C}}, \mathrm{C\%}, \mathrm{I\%}, \mathrm{R\%}\right)$$

$$\left(\frac{\mathrm{WM}}{\mathrm{TS}}\right) = \sum\left(\frac{\mathrm{Pb\ load} + \mathrm{Cu\ load} + \mathrm{Zn\ load}}{\mathrm{TS\ load}}\right) = f\left(\mathrm{ADT\_to}, \mathrm{ADT\_hv}, \frac{\mathrm{V}}{\mathrm{C}}, \mathrm{C\%}, \mathrm{I\%}, \mathrm{R\%}\right)$$

Since the data matrix was small, the equations were validated using Leave-One-Out Cross-Validation (LOOCV) method. The development and cross-validation were performed using MATLAB R2009b (Mathworks Inc, Natick, MA, USA) and the derived MATLAB codes are presented in the Supplementary Information. The validity of the model was analysed using Relative Prediction Error (RPE), Standard Error of Cross-Validation (SECV) and Cross-Validated Coefficient of Determination (Q^2) and the relevant equations are given below (Richardson and Reeves, 2005):

$$RPE = \sqrt{\frac{\sum(Y_{predicted} - Y_{measured})^2}{\sum Y_{measured}^2}} \times 100\%$$

$$SECV = \sqrt{\frac{\sum(Y_{predicted} - Y_{measured})^2}{N}}$$

$$Q^2 = \left[1 - \frac{\sum\left(Y_{predicted} - Y_{measured}\right)^2}{\sum\left(\bar{Y} - Y_{measured}\right)^2}\right] \times 100\%$$

where N is the number of samples, $Y_{predicted}$ is the (FM/TS) or (WM/TS) predicted using MLR equations, $Y_{measured}$ is the (FM/TS) or (WM/TS) calculated using the experimental data and $\bar{Y}$ is the mean of $Y_{measured}$.

The RPE values for FM and WM were 34% and 39%, respectively and these values are considered well within the error limits for a complex system such as an urban road surface (Ni et al. 2001; Mahbub et al. 2011; Egodawatta et al. 2013). Nevertheless, WM has a high SECV (17.5) and low Q^2 (61%) suggesting that the model is less reliable for quantitative prediction, while the Q^2 (86%) and SECV (1.04) values indicate a good fit for FM. Hence, the former should preferably be used only in preliminary investigations, while the latter can be applied for the prediction of FM load. The final equations derived are given below:

$$\left(\frac{FM}{TS}\right) = \sum\left(\frac{\text{Cr load} + \text{Ni load}}{\text{TS load}}\right) = -5.41 - 0.06\,ADT\_to + 0.48\,ADT\_hv + 4.88\,\frac{V}{C} + 1.26\,I\% + 6.90\,R\%)$$

$$\left(\frac{WM}{TS}\right) = \sum\left(\frac{\text{Pb load} + \text{Cu load} + \text{Zn load}}{\text{TS load}}\right) = 32.5 + 0.38\,ADT\_to - 4.22\,ADT\_hv + 36.1\,\frac{V}{C} + 31.2\,I\% + 21.2\,R\%)$$

4. Conclusions

The primary conclusions derived from this study are:

- Mn is likely to be contributed to the build-up by geogenic sources. Pb, Cu and Zn are primarily contributed to road surfaces by brake and tyre wear, Cr and Ni are by vehicle exhaust and Cd is from both exhaust and wear.
- Land use does not exhibit a clear pattern in influencing the metal build-up process, and traffic characteristics primarily influence the metal build-up process. However, it is hypothesised that land use implicitly influences traffic characteristics.
- The prediction equation developed for cumulative metal loads originating from fuel exhaust emissions is relatively more reliable since it has an acceptable level of relative prediction error, low standard error of cross-validation and high cross-validated coefficient.
- Though the relative prediction error is at an acceptable level, the prediction equation developed for cumulative metal loads originating from brake and tyre wear is relatively less reliable and should preferably be used only in preliminary studies because of the high standard error of cross-validation and low cross-validated coefficient.

Supplementary Information

The Supplementary Information provides the methodology used for determining the efficiency of the sample collection procedure, location of study sites and the Matlab codes used for multiple linear regression (MLR) analysis.

5. References

Abdi, H., 2003. Factor rotations, in: Lewis-Beck, M., Bryman, A., Futing, T. (Eds.), Encyclopedia for Research Methods for the Social Sciences. Sage, Thousand Oaks (CA), pp. 978–982.

Adams, M.J. (1995). Chemometrics in analytical spectroscopy. Cambridge [England]: Royal Society of Chemistry.

APHA (American Public Health Association), 2004. Standard Methods for the Examination of Water & Wastewater. Washington DC.

Ball, J.E., Jenks, R., Aubourg, D., 1998. An assessment of the availability of pollutant constituents on road surfaces. Sci. Total Environ. 209, 243–254.

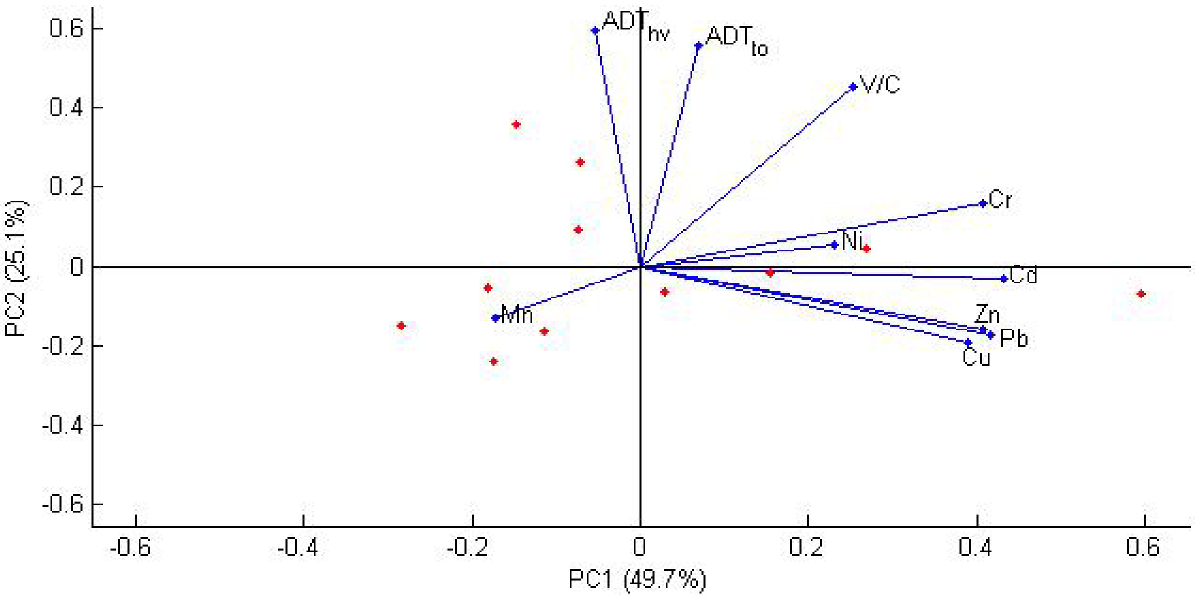
Egodawatta, P., Ziyath, A.M., Goonetilleke, A., 2013. Characterising metal build-up on urban road surfaces. Environ. Poll. 176, 87–91.

EPASGV (Environmental Protection Agency, State Government of Victoria), 1999. Measurement of Motor Vehicle Pollutants and Fleet Average Emission Factors in Melbourne.

Goonetilleke, A., Egodawatta, P., Kitchen, B., 2009. Evaluation of pollutant build-up and wash-off from selected land uses at the Port of Brisbane, Australia. Mar. Poll. Bull. 58, 213–221.

Galvagno, S., Casu, S., Casabianca, T., Calabrese, A., Cornacchia, G., 2002. Pyrolysis process for the treatment of scrap tyres: preliminary experimental results. Waste Manage. 22, 917-923.
Helmreich, B., Hilliges, R., Schriewer, A., Horn, H., 2010. Runoff pollutants of a highly trafficked urban road - Correlation analysis and seasonal influences. Chemosphere 80, 991–997.
Herngren, L., Goonetilleke, A., Ayoko, G.A., 2006. Analysis of heavy metals in road-deposited sediments. Analyt. Chim. Acta 571, 270–278.
Liu, A., Goonetilleke, A., Egodawatta, P., 2012. Inherent errors in pollutant build-up estimation in considering urban land use as a lumped parameter. J. Environ. Qual. 41, 1690–1694.
Mahbub, P., Ayoko, G., Goonetilleke, A., Egodawatta, P., Kokot, S., 2010. Impacts of traffic and rainfall characteristics on heavy metals build-up and wash-off from urban roads. Environ. Sci. Technol. 44, 8904–8910.
Mahbub, P., Goonetilleke, A., Ayoko, G.A., 2011. Prediction model of the buildup of volatile organic compounds on urban roads. Environ. Sci. Technol. 45, 4453–4459.
Mitrović, T., Stamenković, S., Cvetković, V., Đekić, T., Baošić, R., Mutić, J., Anđelković, T., Bojić, A., 2012. Bioindication of heavy metal pollution in the area of Southeastern Serbia by using epiphytic lichen Flavoparmelia caperata (L.) Hale. Biologica Nyssana. 3. 53-60.
Mostert, M.M.R., Ayoko, G. A., Kokot, S., 2010. Application of chemometrics to analysis of soil pollutants. Trends Analyt. Chem. 29, 430–445.
Ni, Y., Wang, L., Kokot, S., 2001. Simultaneous determination of nitrobenzene and nitro-substituted phenols by differential pulse voltammetry and chemometrics. Analyt. Chim. Acta 431, 101–113.
Richardson, A.D., Reeves III, J.B., 2005. Quantitative reflectance spectroscopy as an alternative to traditional wet lab analysis of foliar chemistry: near-infrared and mid-infrared calibrations compared. Can. J. For. Res. 35, 1122–1130.
Sansalone, J.J., Buchberger, S.G., 1997. Partitioning and first flush of metals in urban roadway storm water. J. Environ. Eng. 123, 134–143.
Singh, B., Gilkes, R.J., 1992. Properties and distribution of iron oxides and their association with minor elements in the soils of south-western Australia. J. Soil Sci. 43, 77–98.
US-EPA (US Environmental Protection Agency), 1994. Trace Elements in Waters and Wastes by Inductively Coupled Plasma-Mass Spectrometry. Cincinnati, Ohio.

(a)



(b)

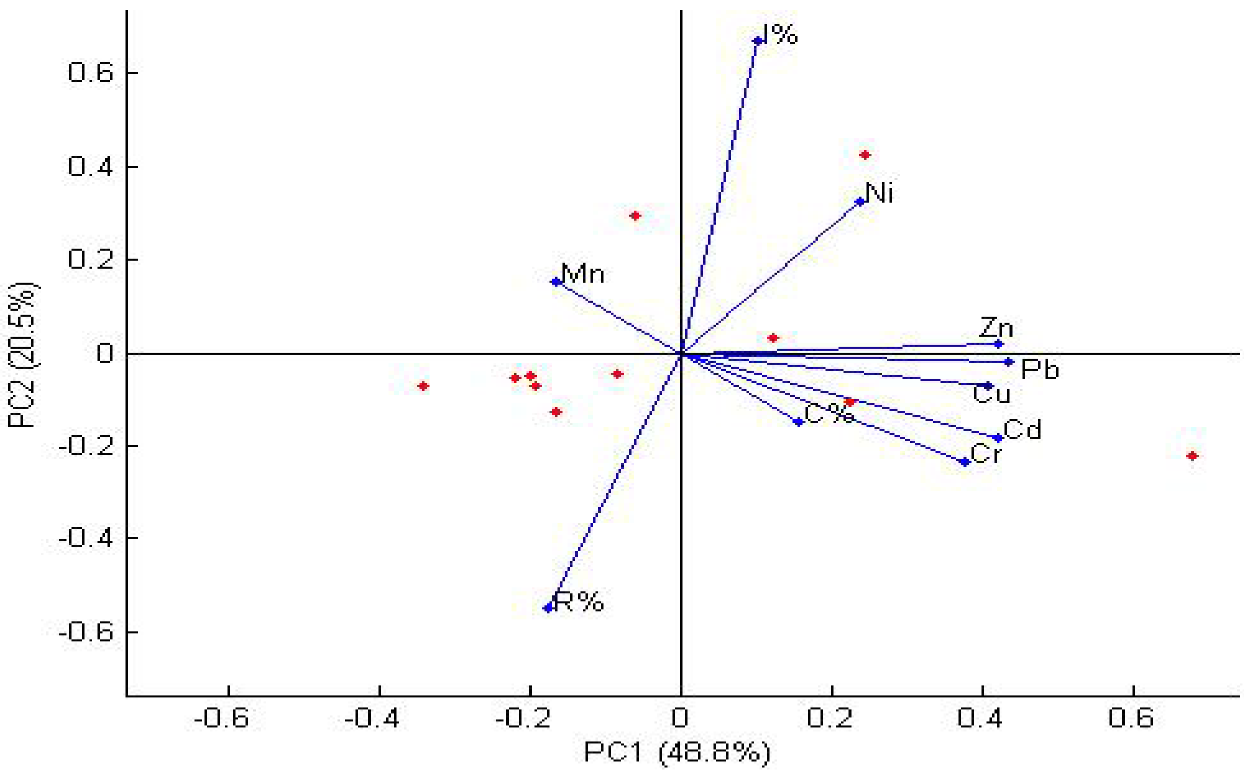


Fig. 1: PCA biplots for: (a) traffic variables (b) land use variables (Legend: Total traffic volume - ADT_to; Total heavy duty traffic volume – ADT_hv; Congestion – V/C; Industrial land use - I%; Residential land use - R%; Commercial land use – C%)

7. Tables

Table 1: Data matrix used in the study

Site ID	Total Solids ($mg/100m^2$)	Metal Loads ($mg/100m^2$)						Traffic Variables			Land use Variables		
		Cd	Cr	Ni	Pb	Zn	Cu	Mn	ADT_to	ADT_hv	V/C	C%	I%	R%
Ab_c	21.2	0.31	2.38	9.11	33.0	233	102	26.5	8739	101	0.60	0.28	0.04	0.68
Re_r	64.4	0.17	19.5	49.1	16.2	256	157	43.2	9973	193	0.72	0.04	0.02	0.94
Pe_r	5.71	0.00	0.00	0.00	12.5	107	75.0	10.6	30	0	0.00	0.03	0.00	0.97
Bi_r	46.9	0.00	0.86	1.24	42.0	351	143	40	1963	10	0.45	0.24	0.03	0.73
Be_i	6.19	0.11	1.11	27.9	32.0	158	90	6.99	4630	86	0.46	0.07	0.59	0.34
Sh_i	18.6	0.00	1.38	1.56	82.9	317	245	39.8	2234	31	0.22	0.05	0.48	0.47
Ho_c	25.2	0.00	2.75	4.90	30.3	247	119	48.2	25571	270	0.59	0.14	0.01	0.85
Li_c	5.13	0.29	2.74	4.25	63.7	209	216	5.47	8594	28	0.73	0.26	0.03	0.71
To_c	6.59	0.05	1.15	7.07	33.4	141	77.3	3.52	5922	61	0.18	0.30	0.17	0.53
Da_r	110	0.00	2.90	5.46	69.3	177	118	190	993	6	0.09	0.01	0.00	0.99
Di_r	6.72	0.19	2.17	21.2	41.7	222	83.4	7.20	10682	41	0.69	0.02	0.00	0.98

Table 2: VARIMAX rotated factor loadings for metal ions

Metals	Factor 1	Factor 2	Factor 3
Cd	0.29	0.11	-0.74
Cr	0.97	0.12	0.01
Ni	0.93	-0.14	-0.18
Pb	-0.39	0.77	0.26
Zn	0.13	0.78	0.00
Cu	0.03	0.94	-0.05
Mn	0.10	0.17	0.89

SUPPLEMENTARY INFORMATION

Mathematical relationships for metal build-up on urban road surfaces based on traffic and land use characteristics

Janaka Gunawardena[a,b], Abdul M. Ziyath[b,c], Prasanna Egodawatta[b], Godwin A. Ayoko[b], Ashantha Goonetilleke[b*]

[a]Toowoomba Regional Council, PO Box 3021, Toowoomba Village Fair, Queensland 4350, Australia

[b]Science and Engineering Faculty, Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia.

[b]Department of Chemical and Process Engineering, Faculty of Engineering, University of Peradeniya, Sri Lanka.

E-mail: mjajanaka@gmail.com; mohamed.ziyath@qut.edu.au; p.egodawatta@qut.edu.au; g.ayoko@qut.edu.au; a.goonetilleke@qut.edu.au

*Corresponding Author: Ashantha Goonetilleke; Tel: 61 7 3138 1539; Fax: 61 7 3138 1170
Email: a.goonetilleke@qut.edu.au

Methodology used to determine the efficiency of the sample collection procedure

100 g of uniformaly graded fine material was spread uniformly on a 1 m x 1 m bituminous surface that was prepared by cleaning with water and dried for an hour. Initially, dry sampling was carried out using the vacuum system followed by wet vacuuming the plots after spraying water at 2 bar pressure for three minutes. The collected sample was carefully transferred to a crucible and oven dried. The total weight recovered through this sampling procedure was calculated and the efficiency of the sample collection procedure was given as the percentage of the initial weight, i.e. 100 g.

Fig. S1: Location map of the study sites (R, C and I represent residential, commercial and industrial land uses).

Matlab codes used for Multiple Linear Regression Analysis

```
function [b,RPE,SECV,Q2] = CVMLR (y,x)
%CVMLR returns multiple linear regression coefficients (b) with relative
%prediction error (RPE), standard error of cross-validation (SECV) and
%cross-validated coefficient of determination (Q2). The input y is the data
%matrix containing dependent variables and x is the data matrix containing
%independent variables.

b = regress (y,x); %Matlab built-in function that performs multiple linear
%regression and returns the regression coefficients matrix, b. The first
%element of matrix b is the constant followed by regression coefficeints.

PredictMatrix = [ ];

for n = 1:length(y)
%Data matrix for model development, consists of all data points except one
   TestIndex = n;
   TrainIndex = setdiff(1:length(y),TestIndex)
   X=x(TrainIndex,:);
   Y=y(TrainIndex,1);

 %Prediction of Y using regression coefficient for the one left out sample
   PredictedY = [x(TestIndex,:)]*b;

 %Data matrix consisting of measured and predicted values
   MeasuredVsPredict = [y(TestIndex,1), PredictedY];
   PredictMatrix = [PredictMatrix; MeasuredVsPredict];
end

[NSamples NVariables] = size(X)
MeasuredY  = PredictMatrix(:,1)
PredictedY = PredictMatrix(:,2)

%Relative error of prediction
RPE = sqrt (sumsqr (PredictedY-MeasuredY)/(sumsqr(MeasuredY)));

% Standard error of cross validation
SECV = sqrt(NSamples/(NSamples-1))*sqrt(sumsqr(MeasuredY-PredictedY)/NSamples);

%Cross-validated R2
Q2 = 1-(sumsqr (PredictedY-MeasuredY)/sumsqr(MeasuredY-mean(MeasuredY)));
```