This is the author's version of a work that was submitted/accepted for publication in the following source:

Egodawatta, Prasanna, Ziyath, Abdul Mohamed, & Goonetilleke, Ashantha
(2013)
Reply to comment on "Characterising metal build-up on urban road surfaces" by Egodawatta et al. (2013).
Environmental Pollution, *182*, pp. 503-506.

This file was downloaded from: http://eprints.qut.edu.au/63302/

Reply to letter to the Editor
Reply to comment on "Characterising metal build-up on urban road surfaces" by Egodawatta et al. (2013). Environmental Pollution, 176, 87–91

Prasanna Egodawatta, Abdul M. Ziyath, Ashantha Goonetilleke
Science and Engineering Faculty, Queensland University of Technology, GPO Box 2434, Brisbane QLD 4000, Australia

Authors response to comments by Zhang and Krebs in relation to the manuscript by Egodawatta, P., Ziyath, A. M. and Goonetilleke, A., 2013, Characterising metal build-up on urban road surfaces, Environmental Pollution, Vol. 176, 87-91.

We welcome the comments by Zhang and Krebs (2013) regarding our recently published manuscript which gives us further opportunity to discuss our research study and its outcomes.

The comments provided by Zhang and Krebs are given in *italics* followed by our responses.

Vacuum sweeper with a water filtration system (heading as per the submission from Zhang and Krebs, 2013)
<u>Comment 1</u>
We have noted the concerns expressed in relation to the sample collection procedure. Our sampling protocol was developed after a thorough review of research literature (for example Bris et al. 1999). Vacuuming was preferred over sweeping and brushing due to its enhanced ability for collecting fine particulates. Water filtration was selected over the typical dust bag collection due to enhanced ability to retain fine particulates.

As we were investigating metal build-up, the samples collected were tested for total metal concentrations. Separating into dissolved and particulate components would not have any meaning. The approach adopted should be evident in the discussion in the Materials and methods chapter and also from the data given in Table S1 in the Supplementary information section.

Therefore, in the context of our study, we do not think that the issue whether the use of a vacuum system will result in the detachment of metals attached to solids is relevant. The sampling procedure was appropriate for the study undertaken.

<u>Comment 2</u>
In paragraph 1 it is stated that,*'Consequently, the pollution potential of RDS was somewhat underestimated by the use of a water filtration system, because a certain proportion of the adhered pollutants will dissolve in the water filtration reagent, and this point was neglected by the authors.'*

The focus of our study was not about assessing the 'pollution potential' of road deposited solids. Consequently, it is not that we have neglected to discuss the 'pollution potential', but rather it is outside the scope of our paper.

<u>Comment 3</u>

In paragraph 2 it is stated that, *'However the RDS samples observed by the authors were started after repeated vacuumings rather than a real rainfall event. Regardless of whether the mistake was derived from a typing error or not, using repeated vacuumings instead of true rainfall to simulate wash-off phenomena will produce further exaggeration and flawed data.'*

We would like to point out that the study focused on pollutant build-up and not pollutant wash-off. Hence, we do not agree that it is necessary to undertake sample collection 'to simulate wash-off phenomena' as it is outside the scope of our study.

<u>Comment 4</u>

In paragraph 2 is it further stated that, *'As stated by the authors, the removal efficiency of repeated vacuumings was found to be 97%, which suggested a much higher removal efficiency than that of a normal rainfall event.'*

The study focus was to investigate the total metal loads present on road surfaces. Therefore, the issue of 'repeated vacuuming resulting in higher removal efficiency than a rainfall event' is not relevant in the context of our study.

<u>Comment 5</u>

In paragraph 2 it is further stated that, *'The initial significantly-higher removal efficiency of repeated vacuumings could further lead to an exaggerated higher build-up rate at the start of antecedent dry weather days (ADDs).'*

We needed to clean the road surfaces as thoroughly as possible in order to assess the total amount of pollutant build-up over the range of antecedent dry periods. We do not see why this would have resulted in exaggerated build-up, but rather the build-up would have been the same as normal circumstances after a cleansing event.

The selection of sampling plots (heading as per the submission from Zhang and Krebs, 2013)

<u>Comment 6</u>

In paragraph 1 it is stated that, *'Each sample collection with respective ADD was performed in sequence, which means that the collected RDS samples were accumulated under different conditions, e.g. weather conditions and traffic conditions, due to the disparate build-up time sequences. This sampling scenario permitted the initial conditions of each specimen to be different, and eventually led to the observed data being unrepresentative and incomparable.'*

After each prescribed antecedent dry period, the road was cleaned as thoroughly as possible using the vacuum system, which enabled pollutant build-up on a clean surface. Therefore, the pollutant accumulation was under actual field conditions for every sampling episode. In the case of rainfall occurrence prior to the completion of the required antecedent dry period, the

road was cleaned once again and the sampling program was repeated. This was to ensure that the initial conditions for each sampling episodes was consistent.

It is not clear what is meant by, 'disparate build-up time sequences' as given above.

It is also not clear what is meant by, 'sampling scenario permitted the initial conditions of each specimen to be different'.

Comment 7
In paragraph 1 it is stated that, *'the surface loads of RDS-adhered HMs with one ADD were in most cases considerably higher than those with 2, 3 or more ADDs. This observation contradicts the widely accepted agreement; i.e. the longer the RDS build-up time, the higher the surface load (Roesner 1982; Sartor and Boyd 1972; Wicke et al. 2012).'*

We totally agree with the outcomes of the research studies cited above. It is evident that this comment appears to be stemming from confusing build-up load with build-up rate. The studies cited above confirm that the build-up rate at 1 ADD is higher than say at 3 ADD and the cumulative build-up load at 1 ADD is lower than the cumulative load at 3ADD. Our study has only re-confirmed this observation and as clearly shown in Fig. 3.

Under these circumstances, the explanations and recommendations provided by Zhang and Krebs (2013) are not deemed to be relevant in the specific context of our study.

Comment 8
Paragraph 2 states that, *'In terms of pollutants build-up efficiency, the largest portion of the sediments was assembled near roadside curb areas (Amato et al. 2010; Sartor and Boyd 1972). Therefore roadside curb areas were widely regarded as a standard sampling location for RDS with regard to temporal accumulation, and are fully discussed by earlier studies (Amato et al. 2010; Sartor and Boyd 1972). However, the middle strip chosen by the authors was not the best choice for sample taking, where the RDS is significantly agitated by traffic flow and was evidently not able to accumulate.*

We agree that the curb area contains the highest pollutant loading. However, this is not representative of the total pollutant loading spread across the entire road surface. The sampling location should be dictated by the nature of the research study being undertaken.

For example, considering the publications cited above, the study by Amato et al. (2010) was to investigate the effectiveness of street sweeping, washing and dust suppressants for removing particulate matter from the urban environment. Therefore, sampling in the region which has the highest particulate loading was the obvious choice.

Similarly, in the case of the seminal study by Sartor and Boyd (1972), it was to investigate pollutant build-up in various regions across a road surface. Once again their conclusion was that the region where the highest pollutant load was present was the curb area. However,

none of these studies have recommended that the curb area should be considered for all research studies in relation to road surface pollution investigations. It was our hypothesis that the middle strip would be representative of the average pollutant loading across the road surface and hence our decision to undertake sampling in this area.

Outlier detection (heading as per the submission from Zhang and Krebs, 2013)
<u>Comment 9</u>
It is stated in paragraph 1 that, *'... some individual data in the groups of P7-2 and G2-3 were not regarded as outliers by the boxplot analysis. In contrast, cluster analysis regarded the whole data sets in P7-2 and G2-3 as outliers, which seems overestimated. Meanwhile, some individual data in the groups of L1-1, L21, and P21 were identified as outliers by the boxplot. However they were ignored by cluster analysis.'*

1. The data in relation to the individual metal species were extracted from the bulk samples collected from the road surface. As the detected outliers could be due to range of potential errors that can interfere with a field program, it was considered important to exclude any results which were obtained using the specific bulk sample that generated an outlier.

2. A boxplot generally determines a data point as an outlier if the data point lies beyond 1.5 or 3 times the interquartile range from both upper and lower quartile of the data matrix. We assume that a similar definition is used by the SPSS (used by Zhang and Krebs, 2013) to determine the outliers. In our analysis, when the factor of 1.5 was used, all data points in G2-3, and some data points in G-21 and P7-2 were found to be outliers. When the factor 3 was used, all points in G2-3 except Cu and some points in P7-2 were found to be outliers. None of the other points listed by Zhang and Krebs (2013) were found to be outliers. In the cluster analysis presented in Egodawatta et al. (2013), we chose only the extreme points as outliers since it is possible in natural systems to have spatial variations. Thus, in line with our approach, we chose a factor of 3 that would account for the extreme case.

<u>Comment 10</u>
It is stated in paragraph 2 that, *'Furthermore, despite the application of cluster analysis for outlier detection being suitable or not, the results obtained by this method focused more on the behavior of group data rather than individual data. However, individual extreme value data play a significant role in the results performed by future factor analysis.*

We believe that our response to Comment 9 above, is also appropriate here.

Factor analysis(heading as per the submission from Zhang and Krebs, 2013)
<u>Comment 11</u>
It is stated that, *'A general issue of concern is one of sufficient sample size when conducting factor analysis, because a desirable level of sample size is essential in achieving stable results. In the present commentary, a subject-to-variables ratio was used to evaluate the sample size of the authors. Although, to date, there is no unified guideline for subject-to-*

variables ratio, some recommendations have been reported in previous studies. …. It is suggested that the data level of the authors is insufficient, and that the reliability of source apportionment is questionable.'

The factor analysis in our study had a subject to variables ratio of 19:7, i.e. 2.7. There are different parameters that can be used to evaluate the reliability of factor analysis such as the sample size (Gorsuch, 1983) and sample to variable ratio (Cattell, 1978). However, as pointed out by Zhang and Krebs (2013), there are no universal guidelines for the acceptable sample size or subject to variable ratio.

Researchers differ widely in terms of acceptable subject to variable ratio. For example, 5:1 (Gorsuch, 1983), 10:1 (Nunnally et al., 1978), 3 to 6:1 (Cattell, 1978) and even 2:1 (Kline, 1979). MacCallum et al. (1999) criticized these recommendations arguing that it is inaccurate to suggest a minimum sample size or minimum subject to variable ratio since the reliability of factor analysis also depends on other parameters such as communality of the variables and over-determination of the factors. MacCallum et al. (2001) found that communality has a significant effect on the reliability of factor analysis such that if it is high, reliable results can still be obtained using factor analysis even for small sample size and weak over-determination.

Furthermore, MacCallum et al. (1999) have suggested that communality for variables greater than 0.6 or a mean communality of at least 0.7 can be considered to be adequate. Based on a detailed study, Preacher and MacCallum (2002) concluded that researchers should not be overly concerned about the rules of thumb relating to sample size and subject to variable ratio if the communalities are high. In this context, the factor analysis in our study is still valid as the communalities for variables were generally high (Al-0.9; Ca-0.4; Cu-0.8; Fe-0.9; Pb-0.4; Mn-0.9; Zn-0.8) and the mean communality was greater than 0.7.

Empirical models (heading as per the submission from Zhang and Krebs, 2013)
<u>Comment 12</u>
It is stated in paragraph 1 and paragraph 2 that, *'The authors used power functions to simulate both the surface load and build-up rate of RDS-associated HMs. The appropriateness of power function simulations need to be discussed in light of the following three points.'*

Note: The paragraph has been separated into sections in order to better respond to the comments made.

1. *'Firstly, in terms of surface load prediction, as reviewed by Huber et al. (1988), a power function may be easily adjusted to resemble asymptotic behavior, but it must always ultimately exceed the maximum value. In this regard the power function proposed by the authors contradicts a widely accepted view that the surface loads level off subsequent to a rapid initial accumulation and gradually approach a limit (Roesner 1982; Sartor and Boyd 1972).'*

We have not proposed a power function (Section 3.3 in our paper). We have said that among the various equations proposed in past research literature, the equation proposed by Ball et al. (1998) was found to perform the best in terms of replicating pollutant build-up. Furthermore, the equation proposed by Ball et al. (1998) is a refinement of the work undertaken by Sartor and Boyd et al. (1972). Based on research that we have undertaken, this is our considered opinion. However, we accept the fact that there could be differing views in this regard.

2. *'In terms of build-up rate prediction, according to the power function proposed by the authors, when the ADD (the independent variable of the power function) approaches zero, the build-up rate (the dependent variable of the power function) approaches infinity. It contradicts the findings of previous studies which reported that the values of build-up rate should follow a certain range (Roesner 1982; Sartor and Boyd 1972; Wicke et al. 2012). As a result, the power functions are not suitable to simulate both the surface load and build-up rate from a mathematical point of view.'*

We are confused by this comment. The equation that we have discussed in the paper which is $B = aD^b$ shows that the build-up rate (B) is directly proportional to antecedent dry days (ADD – D). Not sure how B could approach infinity if D is zero.

Comment 13
It is stated in paragraph 3 that, *'Secondly, from an experimental point of view, the most significant problem with regard to the simulations of the surface load and build-up rate is the unreliable and unrepresentative original RDS build-up data sets used for model calibration and validation. As shown in Fig. 1, the fitness of a power function for the RDS surface load is questionable. Then if the build-up tendency of the metals' surface loads did not occur, we are doubtful of the build-up rate tendency and corresponding models proposed by the authors.'*

We believe we have responded appropriately to the above in our response to Comments 2, 3 and 12.

Comment 14
In paragraph 4 it is stated that, *'Finally, as given in the Table 1 of the author' manuscript, we also question the appropriateness of the proposed surface load models, e.g. the surface load prediction models of Fe, Pb, Cu, and Zn. The time series prediction data of these HMs showed decreasing tendencies, which contradict the widely accepted view of the increase in tendencies as discussed above.'*

We believe we have responded appropriately to the above in our response to Comment 7.

Summary and conclusions(heading as per the submission from Zhang and Krebs, 2013)
Comment 15

It is stated that, *'Although the mechanisms of street surface contaminant build-up were difficult to describe quantitatively, they surely include wind, traffic, rainfall, street sweeping practice, surrounding land use, surface pavement, residence time, season of the year, and so on. In this regard a suitable field sampling protocol is essential to ensure the reliability of the observed data, especially if these data are intended for use in empirical predictive models. The potential uncertainly and inappropriateness induced by the corresponding sampling protocol and data-based predictive models proposed by Egodawatta et al. (2013) need to be reviewed.'*

We most certainly agree that the sampling protocol needs to consider the relevant issues within the context of the research project. The protocol that we adopted was underpinned by appropriate scientific rigor and state-of-the art knowledge.

We do not understand what is meant by the comment by Zhang and Krebs (2013), 'The potential uncertainly and inappropriateness induced by the corresponding sampling protocol and data-based predictive models proposed by Egodawatta et al. (2013) need to be reviewed.' We would like to point out that this paper was published after a rigorous review process as determined by the Journal

References
1. Amato, F., Querol, X., Johansson, C., Nagl, C., Alastuey, A., 2010. A review on the effectiveness of street sweeping, washing and dust suppressants as urban PM control methods. Sci. Total Environ. 408, 3070–3084.
2. 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.
3. Bris, F.J., Garnaud, S., Apperry, N., Gonzalez, A., Mouchel, J. M., Chebbo, G., Thevenot, D.R.,1999. A street deposit sampling method for metal and hydrocarbon contamination assessment. The Science of the Total Environment, 235, 211-220.
4. Cattell, R. B., 1978. The Scientific Use of Factor Analysis. New York: Plenum.
5. Egodawatta P., Ziyath A.M., Goonetilleke A., 2013. Characterising metal build-up on urban road surfaces. Environ. Pollut. 176, 87-91.
6. Gorsuch, R. L., 1983. Factor Analysis. Hillsdale, NJ: Erlbaum.
7. Kline, P., 1979. Psychometrics and Psychology. London: Academic Press.
8. MacCallum, R.C., Widaman, K.F., Zhang, S., Hong, S., 1999. Sample size in factor analysis. Psychol. Meth. 4, 84–99.
9. MacCallum, R.C., Widaman, K.F., Preacher, K., Hong, S., 2001. Sample Size in Factor Analysis : The Role of Model Error. Multivar. Behav. Res. 36, 611–637.
10. Nunnally, J. C., 1978. Psychometric Theory. New York: McGraw-Hill.
11. Preacher, K.J., Maccallum, R.C., 2002. Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. Behav. Genet. 32, 153–161.
12. Roesner L. A., 1982. Quality of urban runoff, in: Kibler D.F. (Ed.), Urban stormwater hydrology. American Geophysical Union, Washington DC, 161-187.
13. Sartor J.D., Boyd G.B., 1972. Water Pollution Aspects of Street Surface Contaminants. US Environmental Protection Agency, Washington DC.
14. Wicke D., Cochrane T.A., O'Sullivan A., 2012. Build-up dynamics of heavy metals deposited on impermeable urban surfaces. Journal of Environmental Management, 113, 347-354.