Development of background soil concentrations for BaP and lead across Newcastle Upon Tyne

Robyne Lowes[a] Martin Cooke[a] Lindsay Bramwell[a, b] Phil Hartley[b]
[a]Newcastle University [b]Newcastle City Council

Introduction

DEFRA has introduced new Statutory Guidance in April 2012[1] due to considerable uncertainty in the determination of concentrations where there is a 'significant possibility of significant harm' (SPOSH).

The new guidance states that concentrations that are 'normal' for an area should not be classed as contaminated[1]. Normal concentrations are defined as either occurring from low level diffuse pollution (from common human activity but not from a specific industry) and from natural geology or soil formation processes resulting in widespread elevated concentrations[1].

Both benzo[a]pyrene (BaP) and lead have been problematic in the past with concentrations in urban areas often above the SGV's which has resulted in large areas of land being remediated at a great cost.

Aim

To develop background soil concentrations for BaP and lead across Newcastle Upon Tyne for rural, urban and industrial areas.

Methods

BaP and lead concentrations from Newcastle City Council's (NCC) existing site investigations were collated into a database, with only soil samples at depths of 0.3m and below recorded. PAH was also added to the database to give an indication of BaP concentrations where they have not been reported. Land has been classified as urban, rural or industrial based on historical land uses, e.g. old mine sites classed as industrial.

The background for the contaminants will be determined using robust statistical methods suggested by Reimann *et al.* (2005):

1.Box and whisker plots: the upper whisker is taken as the background.

2.Median ± 2 median absolute deviation (MAD) is determined as follows:

- The median of the data is identified, and the difference of each value from the median is calculated.
- The median of the calculated differences is the value of MAD
- Background = Median + (2*MAD)

3.Cumulative probability plots: inflections (changes) and breaks in the data (suggesting different populations) is taken as the background value.

For the purpose of the initial investigation samples below the limit of detection (LOD) were given a value of half the LOD[3]. All samples (including outliers) were used in the initial calculations.

Results

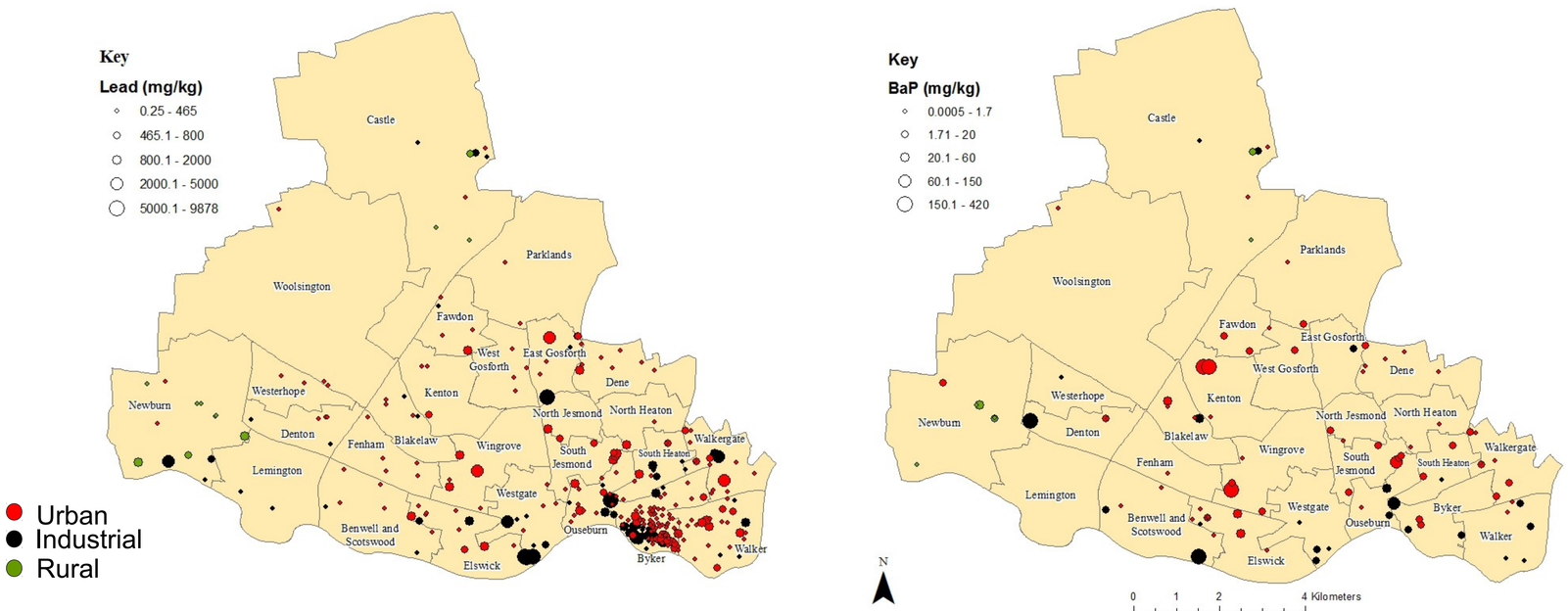

Fig. 1. Lead and BaP soil concentrations across Newcastle Upon Tyne for rural, industrial and urban classes.

Table 1. Summary of results for contaminants in soil

		Min	Max	Median	Mean	Stand. Dev
Lead	Urban	0.25	3200	197	284	304.78
	Industrial	0.5	9878	240	516.74	1099
	Rural	28	1200	159	207.49	218.69
	All	0.25	9878	207	342.71	638.29
BaP	Urban	0.001	170	0.5	2.83	13.16
	Industrial	0.001	420	1	8.61	34.13
	Rural	0.05	46	0.12	1.38	6.9
	All	0.001	420	0.6	4.25	20.74
PAH	Urban	0.05	2500	8.15	52.99	224.45
	Industrial	0.25	6800	19	107.67	446.81
	Rural	0.46	800	2.5	23.85	117.23
	All	0.05	6800	9.576	69.27	312.33

Rural areas have the lowest concentrations for all 3 contaminants, followed by urban, with industrial having the highest (Fig.1). The standard deviations for the results demonstrate the wide range of values, industrial lead for example having a standard deviation of 1099mg/kg (Table 1).

Background Calculation

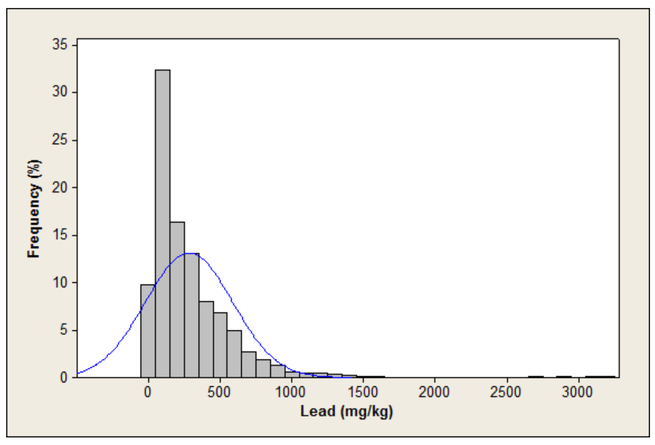

Fig. 2. Histogram displaying right skewed distribution of urban lead soil concentrations.

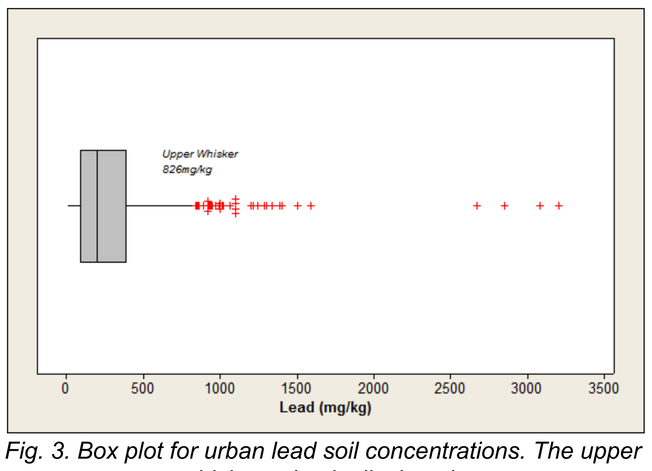

Fig. 3. Box plot for urban lead soil concentrations. The upper whisker value is displayed.

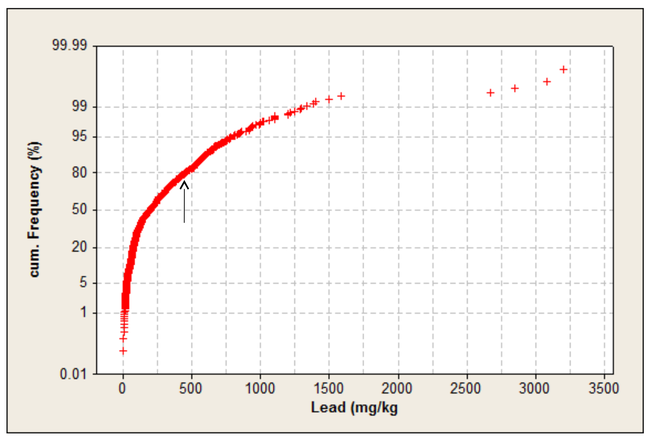

Fig. 4. Cumulative probability plot of urban lead soil concentrations . An inflection is indicated which gives a background of 460mg/kg.

The data distribution for all contaminants is right skewed, as seen from the histogram of urban lead (Fig. 2). The median + 2MAD method gave the lowest background values whereas the box and whisker plot gave a background values of almost double the median + 2MAD for all (excluding industrial and rural lead) (Fig.3, Table 2). The cumulative probability plot method gave the highest values, however this method is subjective, using the eye to spot inflections and breakpoints in the data indicating different populations[2] (Fig 4).

Lead background concentrations are higher than expected, this is due to the inclusion of all samples in the calculations including historically contaminated sites such as old lead works. Therefore once contaminated sites are removed and the background recalculated a lower background for lead may be found which is more representative of the widespread soil concentrations. The BaP concentrations were lower than expected, this may be due to a large percentage of samples <LOD.

BaP and PAH relationship

The correlation between BaP and PAH (Fig. 2) is strong with a Pearson's correlation value of 0.976. Correlations for industrial, urban and rural BaP and PAH are also significant with values of 0.975, 0.981 and 1 respectively. Therefore it may be possible to estimate PAH or BaP concentrations where only one contaminant has been measured.

Table 2. Results of background calculations using normal data and LOD x 0.5.

		Median + 2 MAD	Upper Whisker Box Plot	Cum. Probability plot
Lead	Urban	447	826	460
	Industrial	500	850	740
	Rural	355	445	150
	All	465	810	590
BaP	Urban	1.4	3.7	6.2
	Industrial	2.9	8.8	12
	Rural	0.26	0.91	0.91
	All	1.7	4.6	19
PAH	Urban	19.45	51	163
	Industrial	52	183	220
	Rural	3.4	15	14
	All	23.73	68	196

Conclusion

The results from the initial investigation found the median +2MAD method gave the lowest background values as found by other studies[2]. Elevated lead concentrations are found across Newcastle Upon Tyne with several hotspots of high concentrations often sites of historic lead industries. Calculated BaP background concentrations are above the critical concentration often used as guidance (0.8mg/kg BaP residential soils with plant uptake)[4]. There is a strong relationship between BaP and PAH which can prove useful in estimating values when only one is measured.

Issues:

- Large amount of sample <LOD for BaP and PAH (32% and 41% respectively) and different laboratories have different LOD values.
- Use of existing data – sampling not an equal spread across the entire city, concentrated in certain areas.
- BaP and PAH in many investigations were not both reported.

Next Steps....

- Sensitivity analysis – LOD and the effect it has on the results.
- Data transformation.
- Extreme Outlier removal and recalculation of background.

References

[1]DEFRA (2012) Environmental Protection Act 1990: Part 2A Contaminated Land Statutory Guidance. HM Government.

[2]Reimann, C., Filzmoser, P. And Garrett, R.G. (2005) Background and threshold: critical comparison of methods of determination. *Science of the Total Environment,* 346, p.1-16

[3]Helsel, D.R. (2006) Fabricating data: How substituting values for non detects can ruin results, and what can be done about it. *Chemosphere,* 65, p.2434-2439

[4]Charted Institute of Environmental Health (2009) The CIEH LQM Generic Assessment Criteria for Human Health Risk Assessment (2nd Edition)