DARLINGTON BOROUGH COUNCIL

NORMAL BACKGROUND CONCENTRATIONS OF CONTAMINANTS IN SOILS IN THE BOROUGH OF DARLINGTON

PROJECT REPORT

September 2013

Executive Summary

Following the release of the revised Part 2A Contaminated Land Statutory Guidance in 2012, Defra commissioned the British Geological Survey (BGS) to calculate the normal background concentrations (NBCs) of contaminants in English soils. NBCs are levels of contaminants in soils, which are due to natural and common human anthropogenic processes for a given area. The BGS published the methodology for calculating the NBCs in English soils in 2012.

Darlington Borough Council updated the Contaminated Land Inspection Strategy in January 2013 and stated, if necessary, NBCs will be used as a guide as to what are reasonable levels to support the decision as to whether land within the Borough is contaminated land under Part 2A of the Environmental Protection Act 1990.

This project was undertaken in response to the above and aims to give guidance on the NBC's of contaminants for the Borough of Darlington.

The same seven contaminants selected for the BGS project were selected for this project, to calculate the NBCs, these were: arsenic, benzo[a]pyrene, cadmium, copper, mercury, nickel and lead. Each contaminant was divided into one of two domains: urban and rural.

The calculated NBCs for the Borough of Darlington were based on existing data from previous site investigations undertaken in the Borough. The data used was from greenfield, topsoil samples.

The sample data was subjected to statistical analysis and the value of the upper 95% confidence limit of the 95th percentile was calculated as the value of the NBCs.

The calculated NBCs showed variations between the urban and rural domains for all the 7 contaminants, therefore, indicating that natural and anthropogenic sources influence the values.

The NBCs produced by this project were compared with the NBCs given by the BGS and the SGV/GAC for residential. Variations were found between Darlington Borough Council's NBCs and the NBCs produced by the BGS for English soils. This again indicates that natural and anthropogenic sources influence the values and calculating the NBCs for the local area is beneficial.

The urban NBC for Benzo[a]pyrene was the only contaminant, which was shown to be above the GAC (residential). Darlington Borough Council is confident that the NBCs calculated are typical for the Borough of Darlington and therefore there is no reason to consider them to cause land to qualify as contaminated land (pose an unacceptable risk).

Contents

Executive Summary……………………………………………………………………….. ii
Disclaimer and Acknowledgement..……………………………..……………….......viii
Introduction……………………………………………………………………………..1
Statutory Guidance and Normal Levels……………………………………………...2
Available Contaminant Data…………………………………………………………..4
Methodology……………………………………………………………………………5
Processing samples…………………………………………………………..........5
Statistical Analysis, completed in Excel…………………………………………..8
Normal Background Concentrations for Contaminants………………………………..14
Arsenic (As)…………………………………………………………………………...15
Rural Domain (As)…………………………………………………………………...15
Urban Domain (As)…………………………………………………………………..16
Benzo[a]pyrene (BaP)………………………………………………………………17
Rural Domain (BaP)…………………………………………………………………17
Urban Domain (BaP) ………………………………………………………………..19
Cadmium (Cd)……………………………………………………………………….21
Rural Domain (Cd)…………………………………………………………………..21
Urban Domain (Cd)…………………………………………………………………22
Copper (Cu)………………………………………………………………………….23
Rural Domain (Cu)…………………………………………………………………..23
Urban Domain (Cu)………………………………………………………………….24
Mercury (Hg)…………………………………………………………………………26
Rural Domain (Hg)…………………………………………………………….......26
Urban Domain (Hg)……………………………………………………………......27
Nickel (Ni)…………………………………………………………………………….29
Rural Domain (Ni)……………………………………………………………………29
Urban Domain (Ni)…………………………………………………………………..30
Lead (Pb)……………………………………………………………………………..32
Rural Domain (Pb)…………………………………………………………………..32

Urban Domain (Pb)……………………………………………………………………….34
Limitations of Project ………………………………………………………………...36
Conclusion………………………………………………………………………………...38
References……………………………………………………………………………….39
Appendices……………………………………………………………………………….40
Appendix 1 – Detailed Methodology……………………………………………….40
Appendix 2 – Arsenic Rural Data…………………………………………………..42
Appendix 3 – Arsenic Urban Data……………………………………………………45
Appendix 4 – Benzo[a]pyrene Rural Data…………………………………………...48
Appendix 5 – Benzo[a]pyrene Urban Data…………………………………………..53
Appendix 6 – Cadmium Rural Data…………………………………………………..58
Appendix 7 – Cadmium Urban Data……………………………………………………61
Appendix 8 – Copper Rural Data……………………………………………………..64
Appendix 9 – Copper Urban Data……………………………………………………….68
Appendix 10 – Mercury Rural Data……………………………………………………..71
Appendix 11 – Mercury Urban Data……………………………………………………74
Appendix 12 – Nickel Rural Data………………………………………………………..79
Appendix 13 – Nickel Urban Data………………………………………………………82
Appendix 14 – Lead Rural Data………………………………………………………85
Appendix 15 – Lead Rural Data………………………………………………………90

Figures

Figure 1 – Map showing the urban and rural domain in the Borough of Darlington and the sample point locations 6
Figure 2 – Flow chart showing how to process samples (Johnson et al, 2012).......... 7
Figure 3 – Flow chart for determining the NBC of a contaminant (Johnson et al, 2012) 8
Figure 4 – Edited from Johnson et al, 2012.......... 8
Figure 5 – Edited from Johnson et al, 2012.......... 9
Figure 6 – Edited from Johnson et al, 2012.......... 10
Figure 7 – Edited from Johnson et al, 2012.......... 10
Figure 8 – Edited from Johnson et al, 2012.......... 11
Figure 9 – Edited from Johnson et al, 2012.......... 11
Figure 10 – Edited from Johnson et al, 2012.......... 12
Figure 11 – Normal distribution plot of the untransformed data set for As Rural, showing a Gaussian data set with outliers in the right hand tail 15
Figure 12 – Box Plot for As Rural, shows that the empirical method would not fit well due to outliers.......... 15
Figure 13 – Normal distribution plot for the untransformed data of As Urban, showing a symmetrical data set 16
Figure 14 – Box plot for As Urban 16
Figure 15 – Normal distribution plot for the untransformed data for BaP Rural, shows skewed data 17
Figure 16 – Normal distribution plot for the log transformed data for BaP Rural, still shows a skewed distribution.......... 18
Figure 17 – Normal distribution plot for Box-Cox transformed data for BaP Rural, the distribution is still skewed 18
Figure 18 – Box plot for the original data for BaP Rural, the plot shows that an empirical value is not well defined.......... 18
Figure 19 – Normal distribution plot for the untransformed data set for BaP Urban, showing a skewed distribution.......... 19
Figure 20 – Normal distribution plot for the log transformed data set for BaP Urban, showing a symmetrical distribution.......... 19
Figure 21 – Box plot for the log transformed data for BaP Urban.......... 20
Figure 22 – Normal distribution plot for the untransformed data set for Cd Rural, showing a Gaussian distribution with outliers in the right hand tail.......... 21
Figure 23 – Box plot for Cd Rural 21
Figure 24 – Normal distribution plot for the untransformed data for Cd Urban, showing a symmetrical distribution.......... 22
Figure 25 – Box plot for Cd Urban.......... 22

Figure 26 – Normal distribution plot for the untransformed data set for Cu rural, showing a skewed distribution 23
Figure 27 – Normal distribution plot for the log transform of the Cu Rural data set, showing a Gaussian distribution with outliers in the right hand tail 23
Figure 28 – Box plot of the log transformed data for Cu Rural 24
Figure 29 – Normal distribution plot for the untransformed data set for Cu Urban, showing a symmetrical distribution 24
Figure 30 – Box plot for Cu Urban 25
Figure 31 – Normal distribution plot for the untransformed data set for Hg Rural, showing a Gaussian distribution with outliers in the right hand tail 26
Figure 32 – Box plot for Hg Rural 26
Figure 33 – Normal distribution plot for the untransformed data set for Hg Urban, showing a skewed distribution 27
Figure 34 – Normal distribution plot for the log transformed data for Hg Urban, showing a symmetrical distribution 27
Figure 35 – Box plot for the log transformed data for Hg Urban 28
Figure 36 – Normal distribution plot for the untransformed data for Ni Rural, showing a symmetrical distribution 29
Figure 37 – Box plot for Ni Rural 30
Figure 38 – Normal distribution plot for the untransformed data for Ni Urban, showing a Gaussian distribution with outliers in the right hand tail 30
Figure 39 – Box plot for Ni Urban 31
Figure 40 – Normal distribution plot for the untransformed data set for Pb Rural, showing a skewed distribution 32
Figure 41 – Normal distribution plot for the log transformation of the data set for Pb Rural, still showing a skewed distribution 32
Figure 42 – Normal distribution plot for the Box-Cox transformed data set for Pb Rural, still showing a skewed distribution 33
Figure 43 – Box plot for the original data for Pb Rural 33
Figure 44 – Normal distribution plot for the untransformed data set for Pb Urban, showing a skewed distribution 34
Figure 45 – Normal distribution plot for the log transformed data set for Pb Urban, showing a symmetrical distribution 34
Figure 46 – Box plot for the log transformed data set for Pb Urban 34

Tables

Table 1: 95th percentile values and their confidence limits for As Rural, where L and U are lower and upper percentiles respectively…....……………………………………15
Table 2: 95th percentile values and their confidence limits for As Urban, where L and U are lower and upper percentiles respectively..………………………………………16
Table 3: 95th percentile values and their confidence limits for BaP Rural, where L and U are lower and upper percentiles respectively..…………………………………18
Table 4: 95th percentile values and their confidence limits for BaP Urban, where L and U are lower and upper percentiles respectively..…………………………………20
Table 5: 95th percentile values and their confidence limits for Cd Rural, where L and U are lower and upper percentiles respectively...………………………………………21
Table 6: 95th percentile values and their confidence limits for Cd Urban, where L and U are lower and upper percentiles respectively..……………………………………….22
Table 7: 95th percentile values and their confidence limits for Cu Rural, where L and U are lower and upper percentiles respectively......……………………………………24
Table 8: 95th percentile values and their confidence limits for Cu Urban, where L and U are lower and upper percentiles respectively..……………………………………….25
Table 9: 95th percentile values and their confidence limits for Hg Rural, where L and U are lower and upper percentiles respectively...………………………………………26
Table 10: 95th percentile values and their confidence limits for Hg Urban, where L and U are lower and upper percentiles respectively..………………………………….28
Table 11: 95th percentile values and their confidence limits for Ni Rural, where L and U are lower and upper percentiles respectively…..……………………………………….30
Table 12: 95th percentile values and their confidence limits for Ni Urban, where L and U are lower and upper percentiles respectively..…………………………………31
Table 13: 95th percentile values and their confidence limits for Pb Rural, where L and U are lower and upper percentiles respectively..…………………………………33
Table 14: 95th percentile values and their confidence limits for Pb Urban, where L and U are lower and upper percentiles respectively..………………………………….35
Table 15: Comparing this project with BGS data and GAC/SGV levels.…………….38

Disclaimer and Acknowledgement

Disclaimer

This project is intended to serve as an informative and helpful source of information. Readers must note that legislation, guidance and practical methods are inevitably subject to change and therefore should be aware of current UK policy and best practice. It should be read in conjunction with prevailing legislation and guidance, as amended, whether mentioned here or not. Where legislation and documents are summarised this is for general advice and convenience, and must not be relied upon as a comprehensive or authoritative interpretation. Ultimately it is the responsibility of the person/company involved in the development and/or risk assessment of land to apply up-to-date working practices to determine the contamination status of a site.

Acknowledgment

Darlington Borough Council would like to thank Rachel Binks, a geology student from University of Bristol and resident of Darlington for undertaking the project and writing the report.

Introduction

In response to revised Part 2A Contaminated Land Statutory Guidance in 2012, Defra commissioned the British Geological Survey (BGS) to calculate the normal background concentrations (NBCs) of contaminants in English soils. The research was written by Johnson *et al* (2012) and the results of the project can be found in the report titled “Normal background concentrations (NBCs) of contaminants in English soils: Final Project report”.

The BGS report aimed to give guidance on a broad scale as to the NBCs of contaminants in England and states within the concluding remarks that the methodology used can be applied on a local scale where appropriate data is available.

Under Part 2A Environmental Protection Act 1990, Darlington Borough Council reviewed their Contaminated Land Strategy in January 2013. Darlington Borough Council stated that, if necessary, normal background concentrations (NBCs) will be used as a guide as to what are reasonable levels to support the decision as to whether land within the Borough is contaminated land under Part 2A. This report aims to give guidance on the normal background concentrations of contaminants in the Borough of Darlington.

Normal background concentrations are levels of contaminants in soils, which are due to natural and anthropogenic processes for a given area. For the purpose of this project anthropogenic input is the input from common human activity, such as exhaust fumes and ash released from bonfires, not that from historical and current industrial and commercial uses.

The objectives of this report are:

1) To calculate NBCs for the Borough of Darlington for the same contaminants that the BGS calculated the NBCs for in English soils, in order to be able to compare the results of both studies.

2) To calculate separate NBCs for the urban and rural domains of the Borough of Darlington as natural and anthropogenic processes affecting the two domains will be different.

The aims and objectives of this report will be achieved by following the sample selection guidance and statistical analysis as detailed in the methodology of the BGS report.

Statutory Guidance and Normal Levels

This project has been carried out in order to quantify ‘normal’ levels of contaminants for the Borough of Darlington by calculating the normal background concentrations of contaminants in soils. The definition of ‘normal’ is that as described in the Defra (2012) Statutory Guidance, sections 3.21-3.26:

“3.21 The Part 2A regime was introduced to help identify and deal with land which poses unacceptable levels of risk. It is not intended to apply to land with levels of contaminants in soil that are commonplace and widespread throughout England or parts of it, and for which in the very large majority of cases there is no reason to consider that there is an unacceptable risk.

3.22 Normal levels of contaminants in soil should not be considered to cause land to qualify as contaminated land, unless there is a particular reason to consider otherwise. Therefore, if it is established that land is at or close to normal levels of particular contaminants, it should usually not be considered further in relation to the Part 2A regime and the local authority should have regard to paragraphs 5.2 to 5.4 of this Guidance.

3.23 For the purpose of this Guidance, “normal” levels of contaminants in soil may result from:
(a) The natural presence of contaminants (e.g. caused by soil formation processes and underlying geology) at levels that might reasonably be considered typical in a given area and have not been shown to pose an unacceptable risk to health or the environment.
(b) The presence of contaminants caused by low level diffuse pollution, and common human activity other than specific industrial processes. For example, this would include diffuse pollution caused by historic use of leaded petrol and the presence of benzo[a]pyrene from vehicle exhausts, and the spreading of domestic ash in gardens at levels that might reasonably be considered typical.

3.24 In deciding whether land has normal levels of contaminants, the local authority should consider whether contamination is within the bounds of what might be considered typical or widespread: (a) locally, if there is sufficient information to make a reasonable consideration of what is normal within a local area; and/or (b) regionally or nationally in broadly similar circumstances, having due regard to similarity in terms of land use and other relevant factors such as soil type, hydrogeology, and the form of the contaminants.

3.25 The local authority should decide that normal levels of contaminants exist in relation to land where: (a) those levels are not significantly different to those likely to be typical or widespread within the authority’s area, or in other similar areas; and/or (b) those levels are common or usual in similar land use situations across England or parts of it; and (c) there is no specific reason to consider that those levels of contaminants are likely to pose an unacceptable risk.

3.26 It is possible that specific pieces of land at or slightly above normal levels of contamination with regard to specific substances may pose sufficient risk to be

contaminated land, and that remediation of such land may bring significant net benefits. However, such cases are likely to be very unusual and the authority should take particular care to explain why the decision has been taken, and to ensure that it is supported by robust scientifically-based evidence."
Defra (2012) Statutory Guidance, Sections 3.21-3.26.

Available Contaminant Data

The data for this project was collated from previous site investigations undertaken in the Borough of Darlington. In the majority of cases these were submitted as part of the planning process. These site investigations contained information on a variety of contaminants but information on the seven contaminants analysed in the BGS report (2012) was extracted for analysis. These contaminants were: arsenic (As), benzo[a]pyrene (BaP), cadmium (Cd), copper (Cu), mercury (Hg), nickel (Ni) and lead (Pb).

Topsoil samples reflect both the diffuse anthropogenic input from the atmosphere and land use as well as the natural input from the geology beneath; and as a result are representative samples for this project. Greenfield samples are unaffected by industrial and commercial uses (point sources) and were used in order to give NBCs which would be a reliable benchmark for the background levels of contaminants in the area. No samples containing made ground or that were sampled below the topsoil level as shown on the sample log were included in the data set as they also would not reflect a normal background level. Samples of topsoil from a site derived stockpile for use as a clean cover system were included in the project if they were also greenfield samples. Not all of the selected samples contained contaminant data for all of the seven contaminants being investigated, for example, some were not analysed for Speciated Polycyclic Aromatic Hydrocarbons (PAHs), a group which benzo[a]pyrene is part of.

Methodology

This report is based on the methodology and information supplied by the BGS report (2012), which was commissioned by Defra, but is focused on a much more localised scale, for the Borough of Darlington.

A statistical analysis needed to be performed on the available contaminant data in order to calculate the upper 95% confidence limit for the 95th percentile, the NBC value, of each contaminant. The suitability of each data set needed to be determined, based on the number of samples and their spatial distribution, before an NBC value could be calculated.

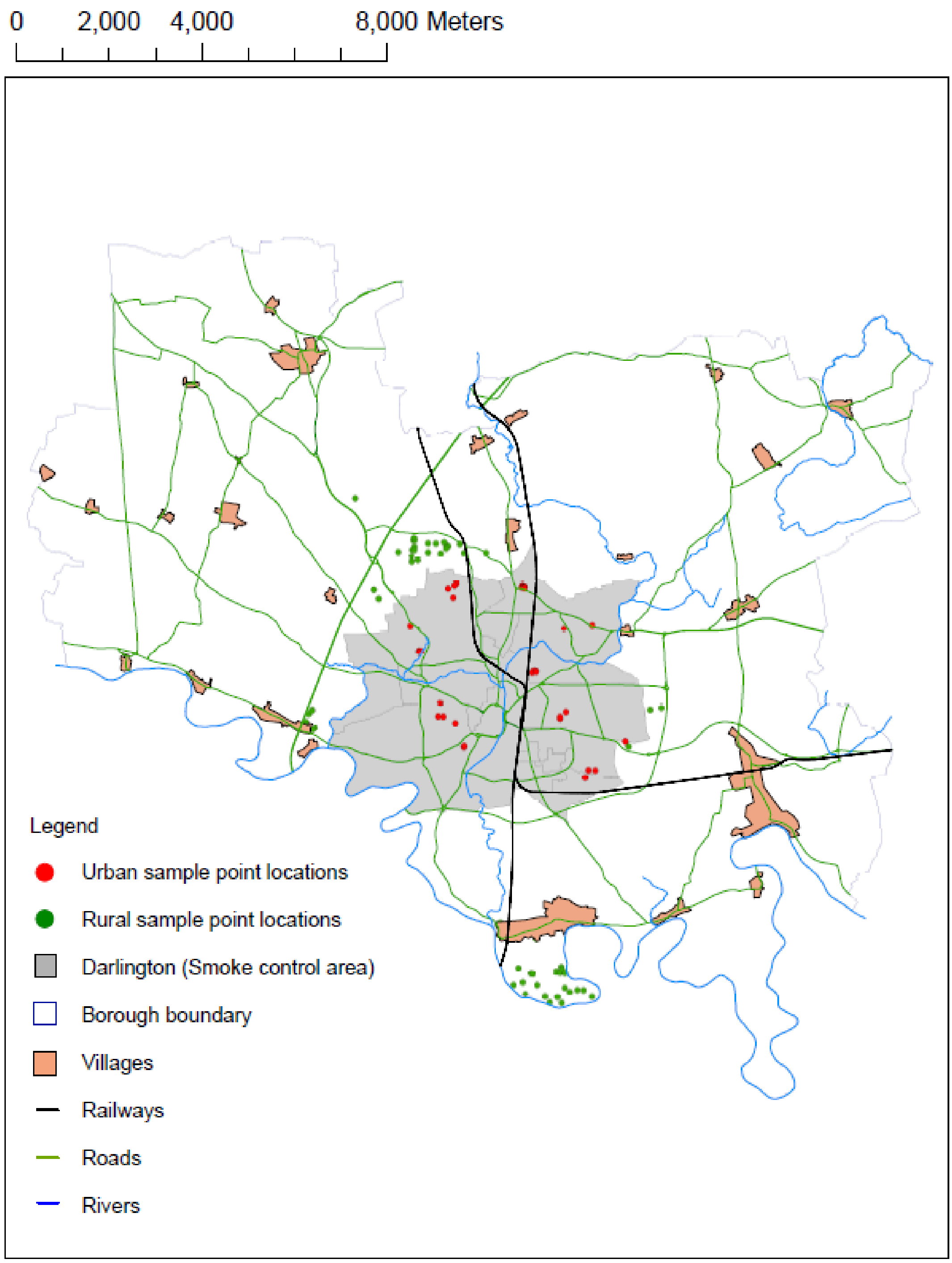
Processing samples

Each contaminant was divided into an urban and a rural domain. The urban domain was classified as within the smoke control area, and the rural domain was classified as outside the smoke control area. The two domains were determined as it was assumed urban and rural levels of contaminants were likely to differ from each other. The spatial distribution of samples across urban areas was good, it was poorer in rural areas, but it was decided that the data spread was adequate to continue with the analysis.

In Figure 1 rural and urban sample domains are differentiated between by coloured dots. Rural samples are denoted by a green dot and urban samples by a red dot.

As the spatial distribution of the data was considered appropriate, each of the 7 contaminants to be analysed were divided into the two domains, urban or rural. This gave 14 data sets for use in the statistical analysis, these varied in size depending on how many samples contained data on each particular contaminant.

The methodology produced by Defra and BGS (2012) recommends no less than 30 results should be used to calculate normal background levels so the total number of results in each data sheet had to be checked. The minimum number of available sample results was 30, and the majority of data sets consisted of a significantly higher number of sample results, so all data sets were appropriate for use in the statistical analysis. This meant no further sampling work needed to be undertaken for this project. The plots produced of the data appear to consist of fewer sample points than they actually contain due to the stacking of points of the same value.



DARLINGTON
BOROUGH COUNCIL

N

Figure 1 – Map showing the urban and rural domain in the Borough of Darlington and the sample point locations

Figure 2 below shows how to process samples according to the BGS methodology (2012):

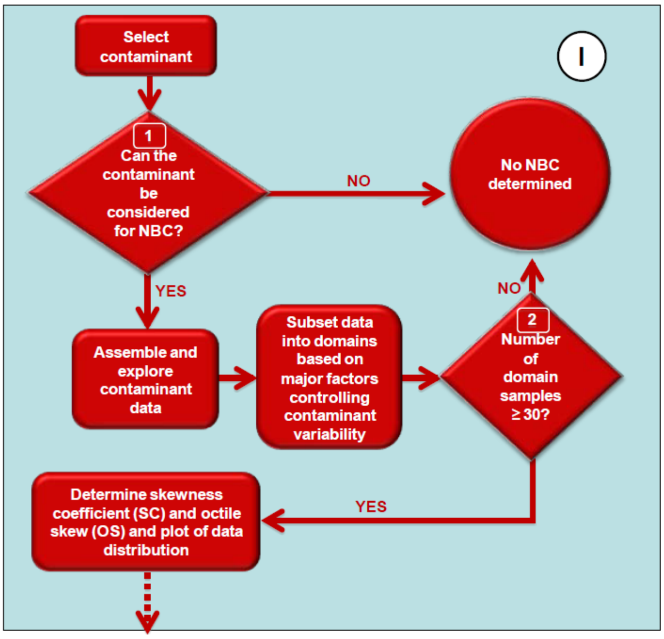


Figure 2 – Flow chart showing how to process samples (Johnson et al, 2012)

Following the sample processing as shown in the flow chart in Figure 2 there was confidence that the data sets would provide appropriate guidelines for NBCs for the Borough of Darlington.

Statistical Analysis, completed in Excel

The methodology produced by the BGS also contains a flow chart for determining the NBC of a contaminant:

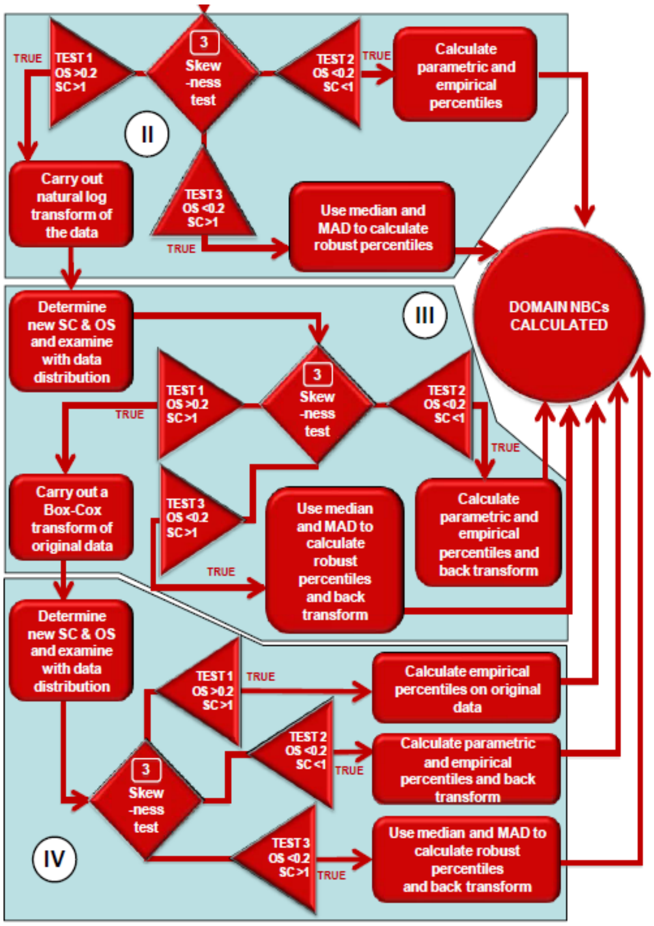


Figure 3 – Flow chart for determining the NBC of a contaminant (Johnson et al, 2012)

This project followed the process shown in Figure 3 in order to calculate the NBCs of contaminants for the Borough of Darlington.

The first stage of the statistical analysis was to determine the type of distribution that the data set displays. This was the skewness test portion of the flowchart, as shown in Figure 4.

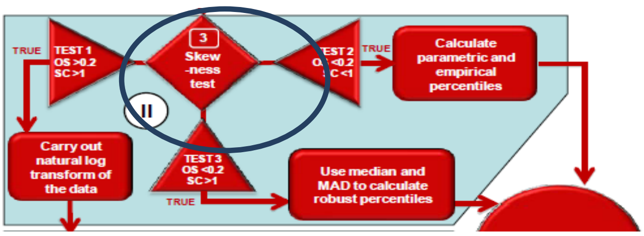


Figure 4 – Edited from Johnson et al, 2012

A normal distribution plot was created in Excel in order to see a visual representation of the data set. The shape of the normal distribution plot can show if variation within the data set is due to normal random variation or if it is due to one specific point source, i.e. sample result, which was taken into account when calculating the NBCs.

To get a statistical definition of the shape of the data sets the skewness coefficient (SC) and octile skewness coefficient (OS) were calculated.

The skewness coefficient is calculated as:

$$SC = \sum(x^i-\mu)^3/N\sigma^3$$

Where μ is the mean, σ is the standard deviation and N is the number of data points

The octile skewness co-efficient is calculated as:

$$OS = ((Q_{0.875}–Q_{0.5})–(Q_{0.5}–Q_{0.125}))/(Q_{0.875}–Q_{0.125})$$

Where Q_n is the n^{th} quantile of the data set

The SC and OS values were used to classify the shape of the distribution, and determine whether the transformation processes shown in Figure 5 needed to be applied to the data set. Transformation was required if a data set gave a Test 1 result from the skewness test. Transformation of the data set attempts to produce an approximately Gaussian distribution so that percentiles can be fitted to it.

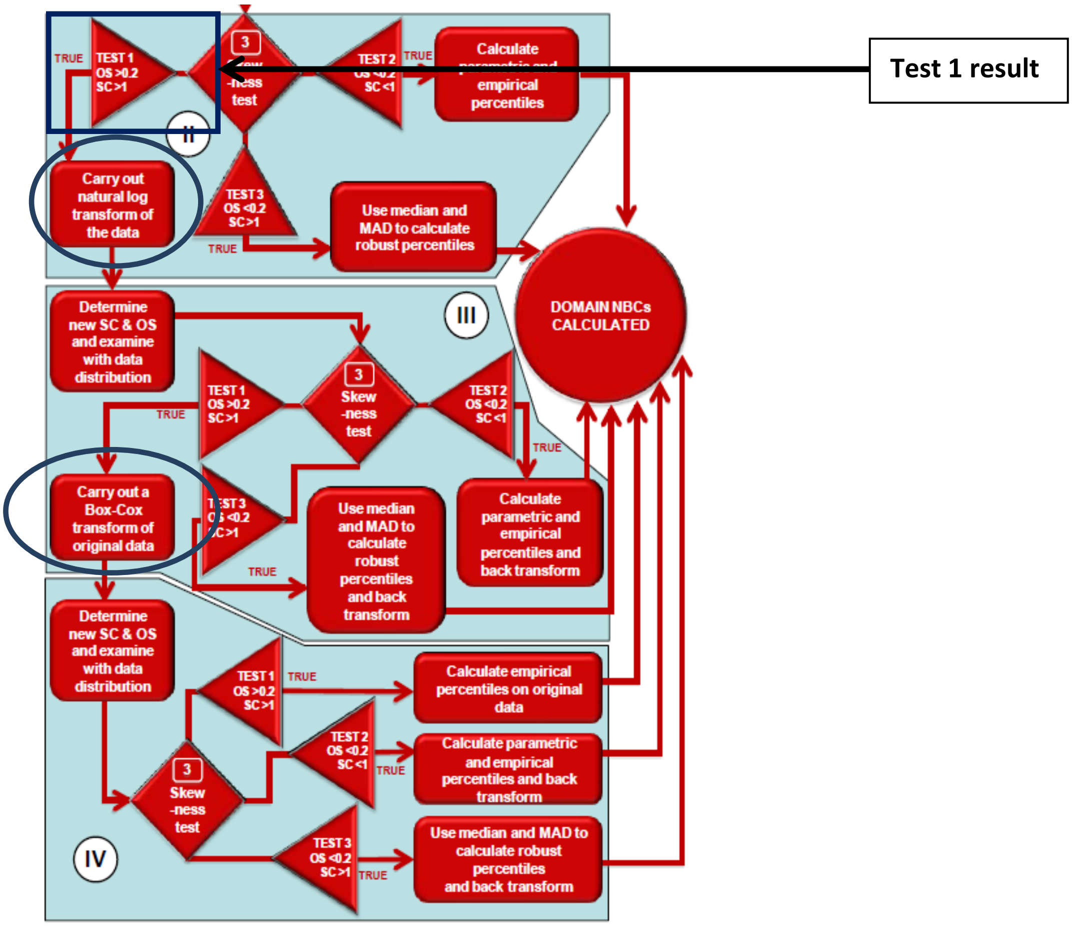


Figure 5 – Edited from Johnson et al, 2012

The data sets could be classified into three different shapes, as shown in Figure 6.

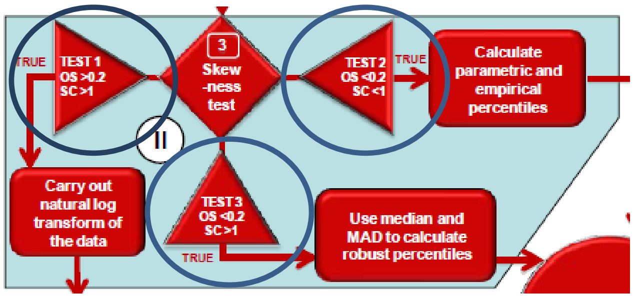


Figure 6 – Edited from Johnson et al, 2012

1) Test 2, symmetrical distribution, SC<1 and OS<0.2
2) Test 3, Gaussian distribution with outliers in the right hand tail, SC>1 but OS<0.2
3) Test 1, non-Gaussian, a skewed data set, SC>1 and OS>0.2

The processing of the data set depended on its classification; Figure 7 shows the first section of the flow chart. A symmetrical distribution, Test 2 on Figure 7, needed no transformations applying to the data set, so percentiles were calculated and the NBC calculated. A Gaussian distribution with outliers in the right hand tail, Test 3 on Figure 7 also did not require any further processing, robust percentiles were fitted to the data and the NBC calculated. A non-Gaussian distribution, Test 1 on figure 7, required a transformation of the data set.

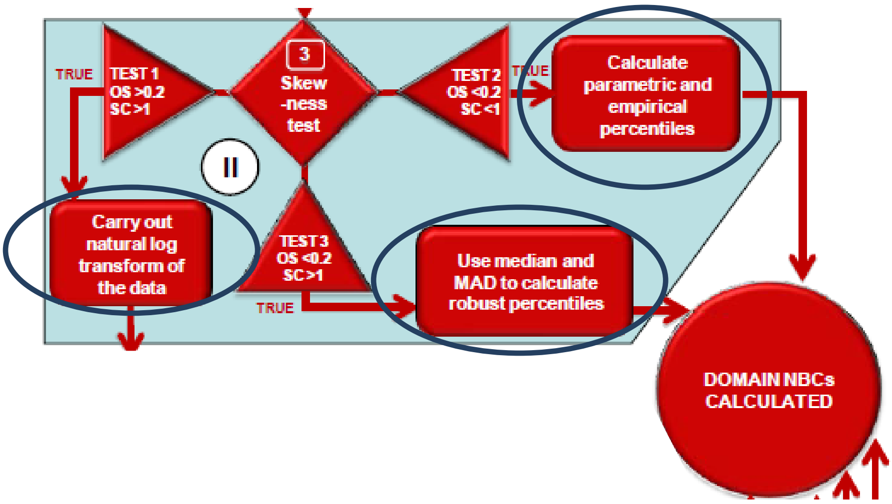


Figure 7 – Edited from Johnson et al, 2012

Following the pathways shown in Figure 7, a non-Gaussian data set was subjected to a log transformation. The shape of the data set was determined by repeating the skewness tests, as shown in Figure 8. The shape of the data set doesn't change whether natural logs or log to the base 10 is used, so the data sets are unaffected by the choice made for this project.

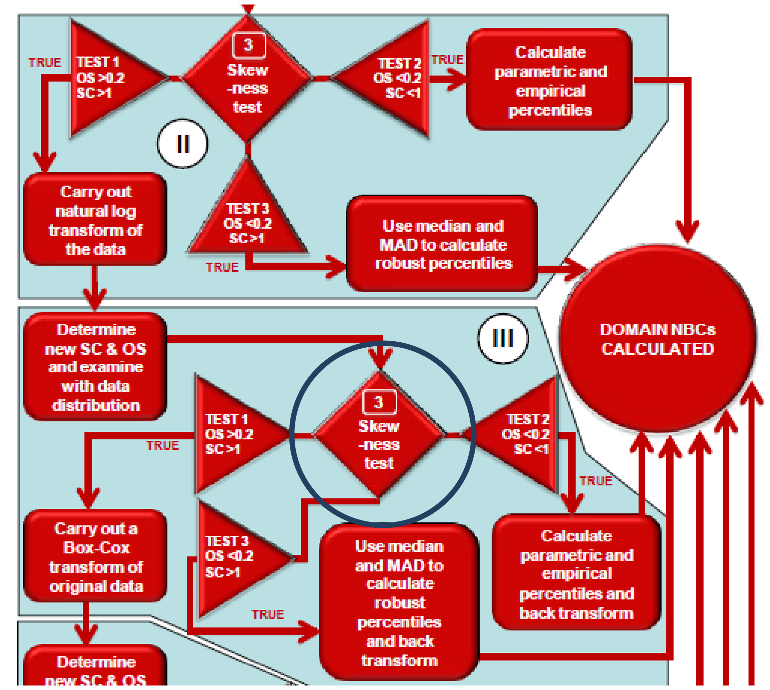


Figure 8 – Edited from Johnson et al, 2012

The skewness test was repeated on the new data set, and then the data set was classified again. The same pathways were followed as in Figure 7. If a data set was classified as Test 2 or Test 3 percentiles were fitted onto the log transformed data set. If a data set came out as Test 1 then the data set required further transformation.

If another transformation was required the original data was put through the Box-Cox transformation, as shown in Figure 9.

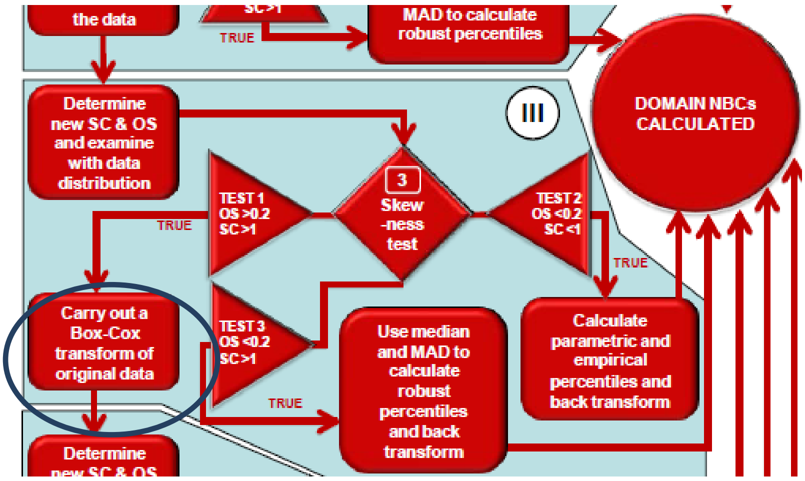


Figure 9 – Edited from Johnson et al, 2012

A statistical add-in for Excel, QI Macros, was used to calculate the Box-Cox transformation. The resulting data set was then subjected to the skewness tests and classification again, as shown in Figure 10.

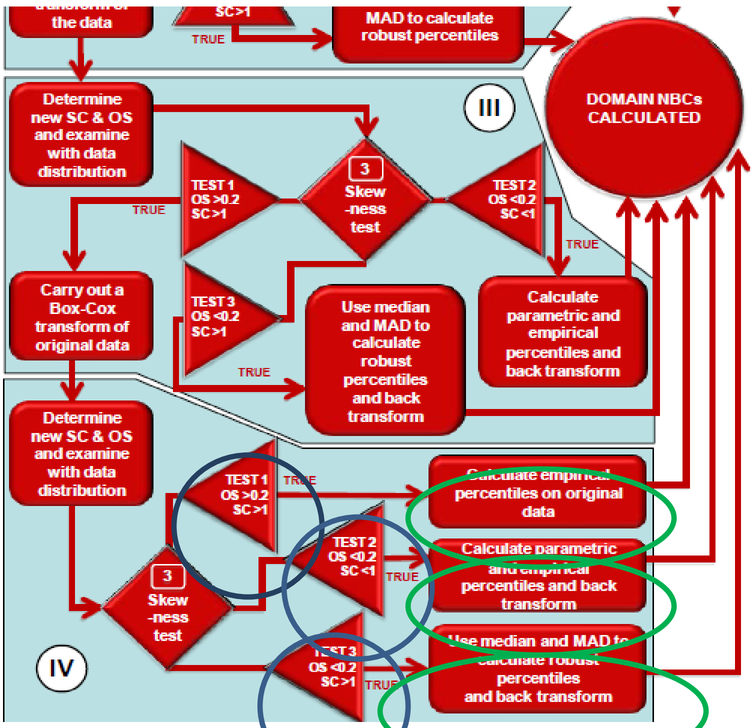


Figure 10 – Edited from Johnson et al, 2012

If the transformed data set had a Test 2 or Test 3 result then percentiles could be fitted. If a data set showed a Test 1, non-Gaussian, distribution no more transformations could be applied and the percentiles were fitted to the original data set.

Three different methods for calculating percentiles were used for the data sets. Empirical, parametric and robust percentiles were calculated. Empirical and parametric percentiles were fitted to the data set using the mean and standard deviation. The robust method for percentiles used the median and MAD of the data set instead and the 95th percentile was determined using median±2MAD.

All three methods were applied to all data sets to get a comprehensive set of data. By following the flow chart in Figure 3, as shown above, the appropriate percentile method for each data set was determined.

For two of the data sets in this project, benzo[a]pyrene rural and lead rural, the distribution could not be approximated as Gaussian even after the Box-Cox transformation had been applied. The Defra, BGS report (2012) recommended in this case that empirical percentiles should be fitted to the original data to calculate the NBC. Based on the results of this project and how the empirical percentiles compared with the other two methods a decision was made to go against this methodology. Instead it was decided to fit robust percentiles to the original data sets, full reasoning for this decision can be found in the sections relating to the rural domains of benzo[a]pyrene and lead, later in the report.

As the actual value of the NBC is defined as “the upper 95% confidence limit of the 95th percentile” upper and lower confidence limits were fitted to the 95th percentile of a contaminant for all three methods. The upper 95% confidence limit for the 95th percentile, determined by the appropriate method, for each data set was taken as the value of the urban or rural NBC for the Borough of Darlington.

Refer to Appendix 1 for more detail regarding the methodology followed for this project.

Normal Background Concentrations for Contaminants

The Borough of Darlington encompasses the large town of Darlington and the surrounding villages. The town of Darlington falls within the smoke control area of the Borough, so is the urban domain. The villages and land around the town of Darlington that fall outside the smoke control area, is the rural domain.

Darlington has an industrial past, due to the town's involvement with the railways. Small foundries producing iron and steel evolved and developed in response to the growing railway network, until large iron-making companies were established. When the foundries eventually went into decline, heavy engineering work took over in Darlington, forging metals for wagon workings and the railways. Point sources of contamination from the localised inputs of these industries will not have affected the data sets in this project because of the sample selection process.

The urban domain of the Borough has a much higher volume of people and traffic than the rural domain, so the anthropogenic input due to exhaust fumes and other human activities is higher in the urban domain. The main anthropogenic input in the rural domain is from agricultural practices such as the use of fertilisers and pesticides.

The solid geology of the Borough is predominantly made up of Magnesian Limestone, with areas of Millstone Grit, Triassic Sandstone, Permian Mudstones and Coal Measures. The drift geology atop this consists mainly of glacial till, with some areas of river terrace and alluvium.

Arsenic (As)

Arsenic is a metalloid that can occur naturally in the environment from the weathering of rocks, but also anthropogenic sources, such as pesticides. High concentrations of arsenic can come from the weathering of shales and mudstones in particular. Igneous rocks also contain arsenic but not in as high concentrations.

The NBC of arsenic was not expected to be high in the Borough of Darlington given the natural and anthropogenic sources, the results below agree with this.

Rural Domain (As)

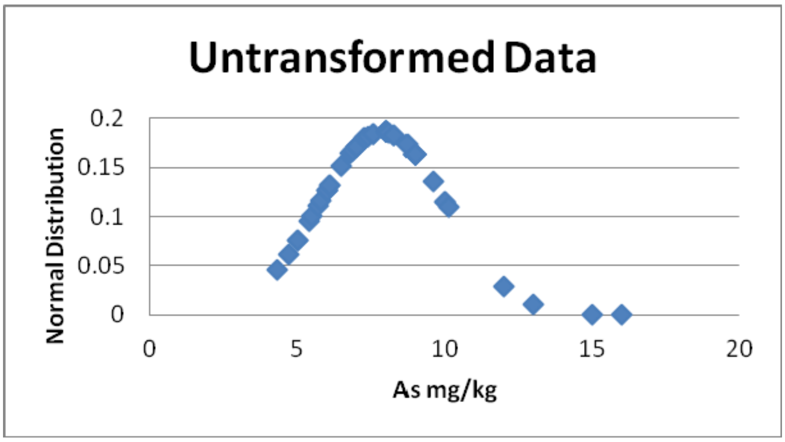


Figure 11 – Normal distribution plot of the untransformed data set for As Rural, showing a Gaussian data set with outliers in the right hand tail

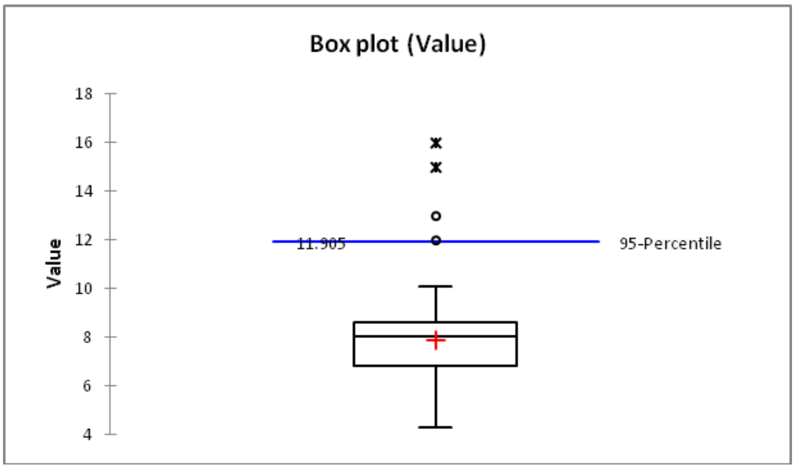


Figure 12 – Box Plot for As Rural, shows that the empirical method would not fit well due to outliers

Table 1: 95th percentile values and their confidence limits for As Rural, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp U	Parametric	P L	P U	Robust	R L	**R U**
61	11.905	10	16	12	11.46	12.54	8.23	8.20	8.25

Robust percentiles are used to calculate the NBC for this data set because it shows a Gaussian distribution with outliers in the right-hand tail. See Appendix 2 for further information.

The NBC for the arsenic rural domain is 8.3mg/kg.

Urban Domain (As)

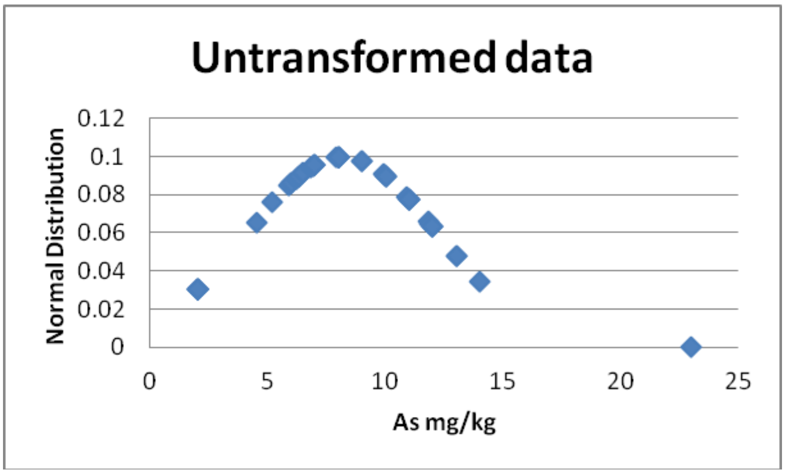


Figure 13 – Normal distribution plot for the untransformed data of As Urban, showing a symmetrical data set

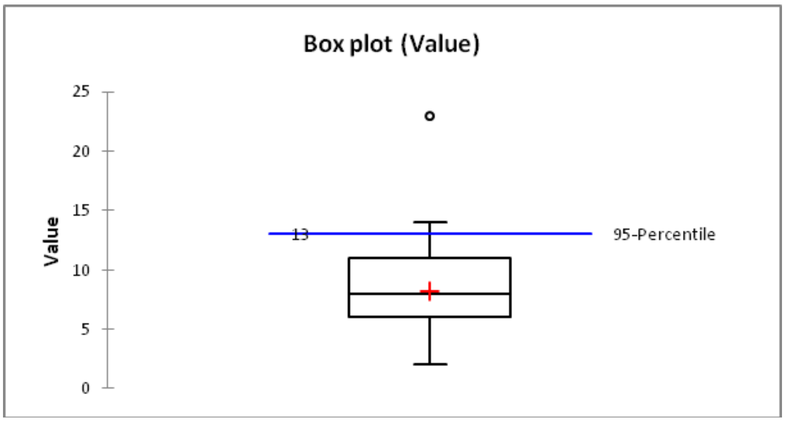


Figure 14 – Box plot for As Urban

Table 2: 95th percentile values and their confidence limits for As Urban, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	R H
48	13	12	23	13	11.87	14.13	8.33	8.28	8.38

Parametric percentiles are fitted to this data set because it shows a symmetrical distribution. See Appendix 3 for further information.

The NBC for the arsenic urban domain is 14.1mg/kg.

Benzo[a]pyrene (BaP)

Benzo[a]pyrene is a Polycyclic Aromatic Hydrocarbon (PAH), which tends to be found in the environment due to incomplete combustion of carbon-rich materials. It has both natural and anthropogenic sources. Natural sources are typically volcanic eruptions and forest fires, whereas anthropogenic sources tend to be from the burning of fossil fuels and wood, exhaust fumes, the use of creosote, combustion of refuse, garden bonfires, the spreading of domestic ash in gardens and also from tobacco smoke.

There are more anthropogenic than natural sources so levels of benzo[a]pyrene are more directly affected by human activity, as a result it was expected that urban levels would be significantly higher than rural levels. (GAC report, Nathanail *et al*, 2009)

Benzo[a]pyrene levels in Darlington were considered likely to be elevated because of the urban infrastructure of the town. Darlington has a high volume of traffic which could lead to increased levels of benzo[a]pyrene from exhaust fumes. The large number of people in the urban area of Darlington was also considered likely to increase levels of benzo[a]pyrene because of the high level of human activity.

The main source of benzo[a]pyrene in Darlington was therefore expected to be anthropogenic so a significant difference between urban and rural areas was predicted. This was in keeping with the GAC report, Nathanail *et al*, 2009, which found that urban soils tend to have 5-8 times higher levels than rural soils.

The results below agree with these predictions. Benzo[a]pyrene is significantly higher in urban areas of Darlington than in rural and the NBC for Darlington is higher than the GAC level.

Rural Domain (BaP)

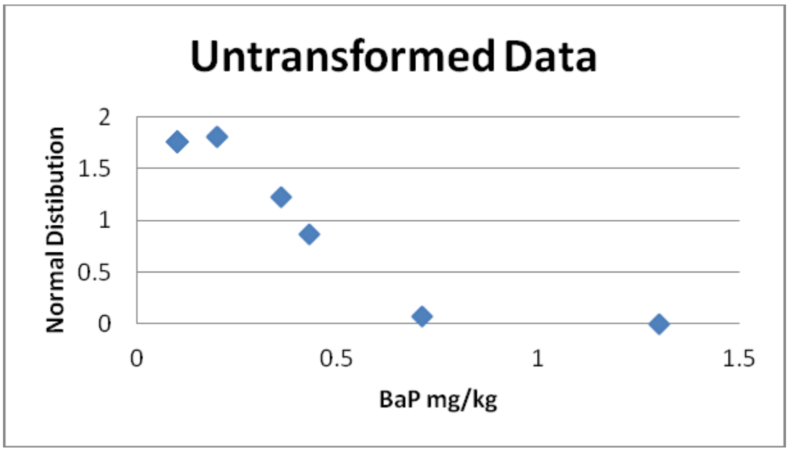


Figure 15 – Normal distribution plot for the untransformed data for BaP Rural, shows skewed data

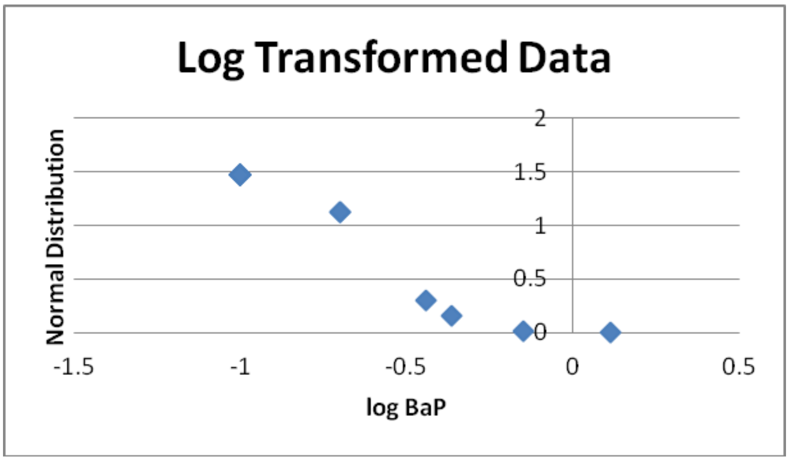


Figure 16 – Normal distribution plot for the log transformed data for BaP Rural, still shows a skewed distribution

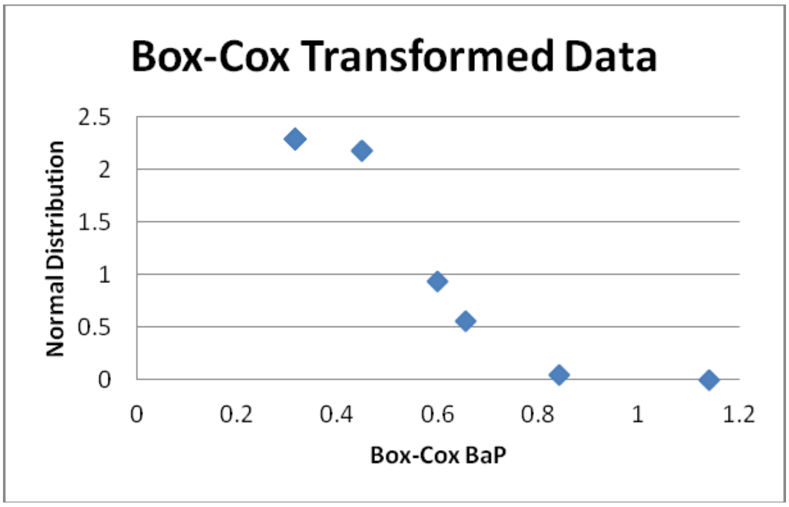


Figure 17 – Normal distribution plot for Box-Cox transformed data for BaP Rural, the distribution is still skewed

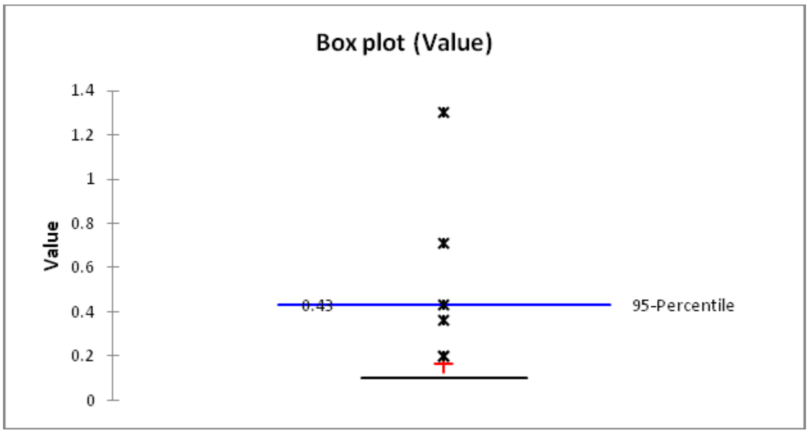


Figure 18 – Box plot for the original data for BaP Rural, the plot shows that an empirical value is not well defined

Table 3: 95th percentile values and their confidence limits for BaP Rural, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	**R H**
40	0.43	0.2	1.3	0.44	0.38	0.51	0.23	0.21	**0.25**

This data set did not show a normal distribution which could approximate as Gaussian even after transformations of the original data. The BGS guidelines suggest returning to the original data set in this case and fitting empirical percentiles. Using the XLSTAT program to fit the empirical percentiles along with their confidence limits gave the 95% confidence limit of the 95th percentile, so the NBC, as the maximum value of the data set. This does not give a good representation of the data set. After contemplation it was decided for this project to instead use the robust percentile method to fit the NBC for benzo[α]pyrene. This decision was reached because the normal distribution plot for benzo[a]pyrene shows half of a bell curve that had been significantly skewed to the right hand side. The robust percentile has been fitted to complete bell curves in this project under the BGS guidelines as this method is robust to outliers, so it was decided it could be used for this data set too in order to give an appropriate value for the NBC. See Appendix 4 for futher information.

The NBC for the benzo[a]pyrene rural domain is 0.25mg/kg.

Urban Domain (BaP)

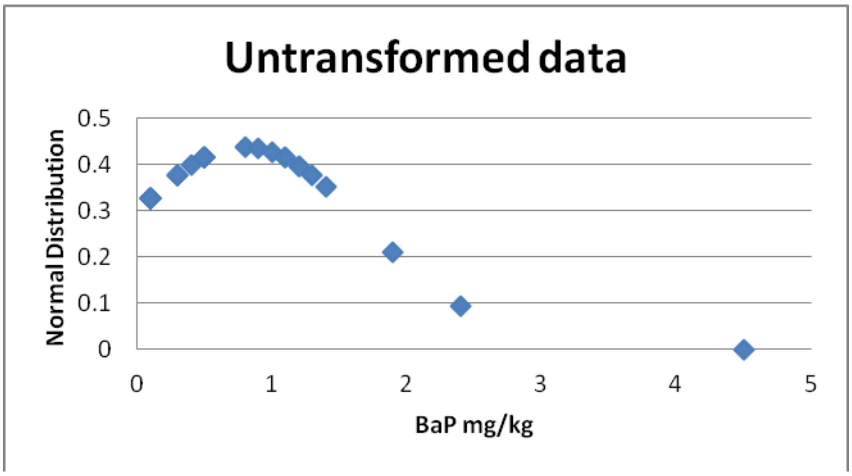


Figure 19 – Normal distribution plot for the untransformed data set for BaP Urban, showing a skewed distribution

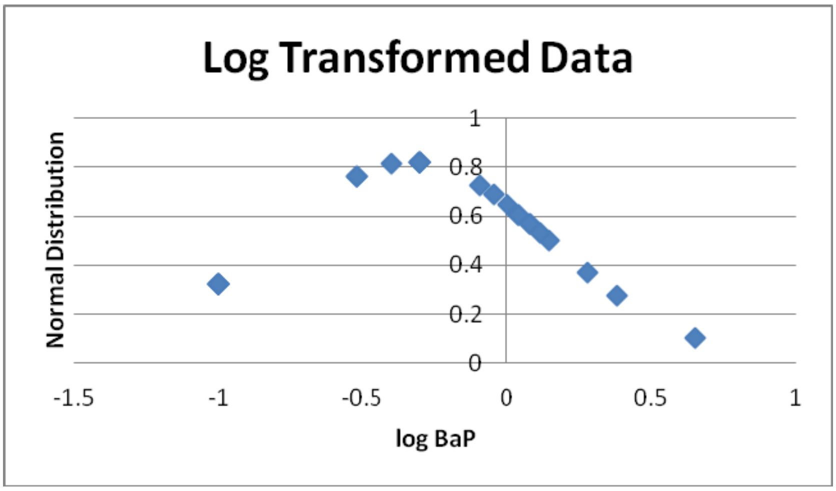


Figure 20 – Normal distribution plot for the log transformed data set for BaP Urban, showing a symmetrical distribution

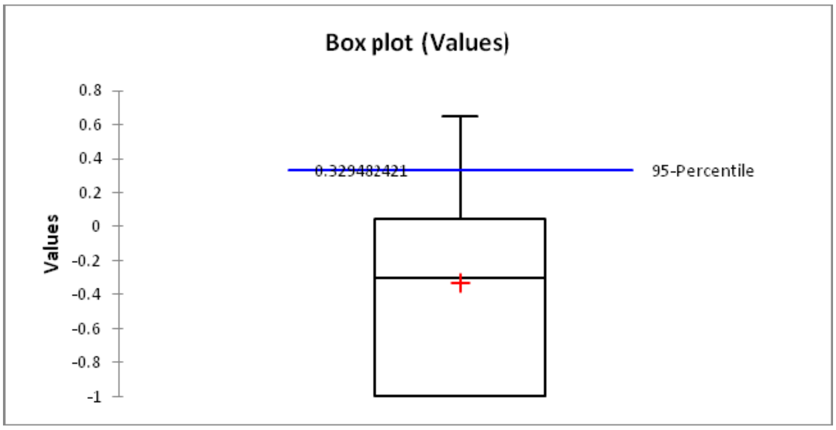


Figure 21 – Box plot for the log transformed data for BaP Urban

Table 4: 95th percentile values and their confidence limits for BaP Urban, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	**P H**	Robust	R L	R H
30	2.14	1.3	4.5	2.16	1.45	**3.23**	0.59	0.57	0.61

Parametric percentiles were fitted to the log transform of this data set as it showed a symmetrical distribution. Once the percentiles had been fitted using the transformed data, inverse logs were performed on the percentiles to give the actual value of the NBC. See Appendix 5 for further information.

The NBC for the benzo[a]pyrene urban domain is 3.2mg/kg.

Cadmium (Cd)

Cadmium is a metallic element, which is found in batteries, plastics, glasses, PVC stabilisers, protective platings on steel and in various alloys. Natural sources of cadmium tend to be from volcanic activity and forest fires. Anthropogenic sources of cadmium include the inappropriate disposal of waste, fertilisers and coal combustion. (GAC report, Nathanail *et al*, 2009)

The NBC of cadmium was not expected to be high in the Borough of Darlington or show a significant variation between the urban and rural domain given the natural and anthropogenic sources, the results below agree with this.

Rural Domain (Cd)

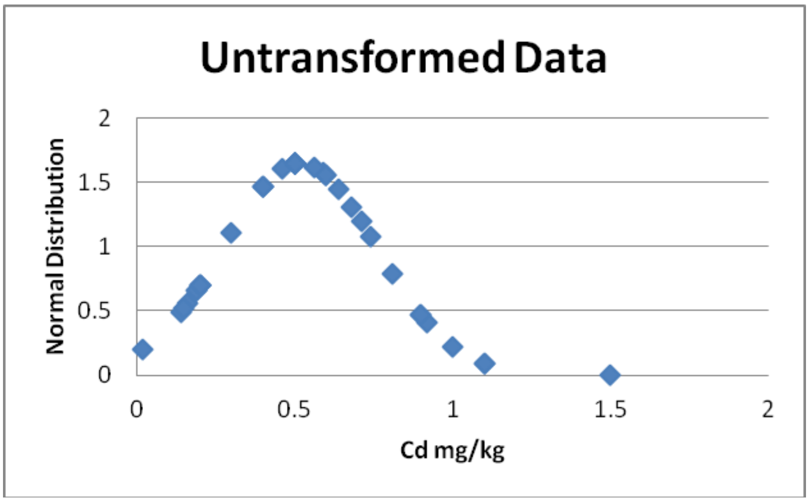


Figure 22 – Normal distribution plot for the untransformed data set for Cd Rural, showing a Gaussian distribution with outliers in the right hand tail

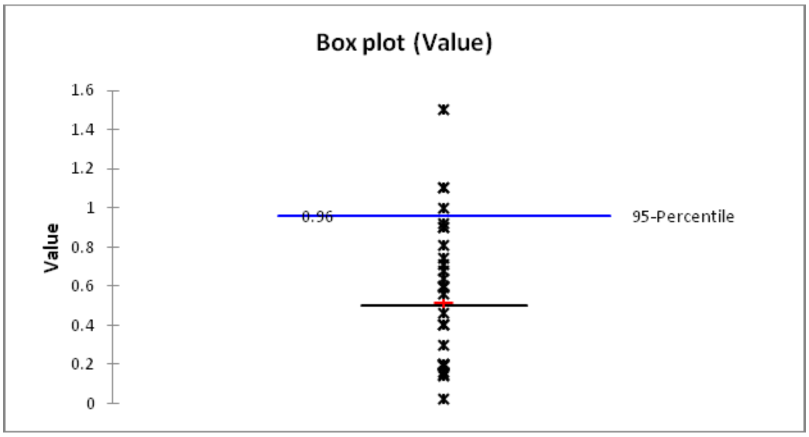


Figure 23 – Box plot for Cd Rural

Table 5: 95th percentile values and their confidence limits for Cd Rural, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	**R H**

70	0.96	0.81	1.5	0.96	0.91	1.02	0.53	0.53	**0.54**

Robust percentiles were fitted to this data set because it showed a Gaussian distribution with outliers in the right hand tail. See Appendix 6 for further information.

The NBC for the cadmium rural domain is 0.54mg/kg.

Urban Domain (Cd)

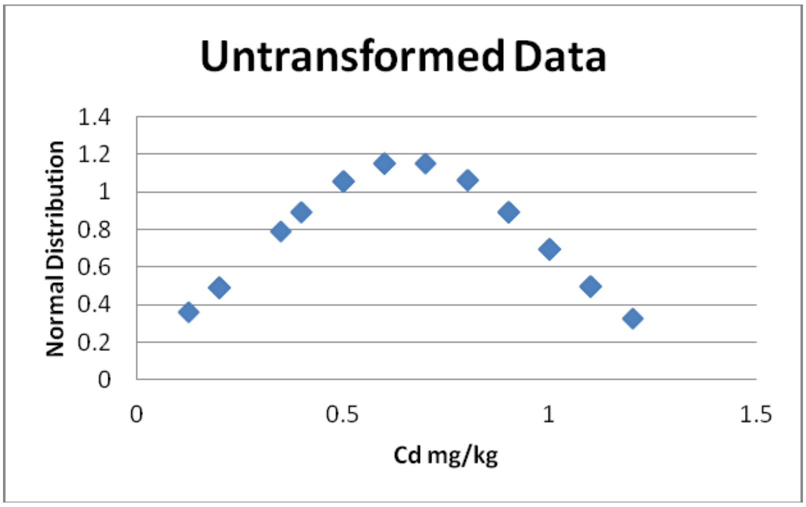


Figure 24 – Normal distribution plot for the untransformed data for Cd Urban, showing a symmetrical distribution

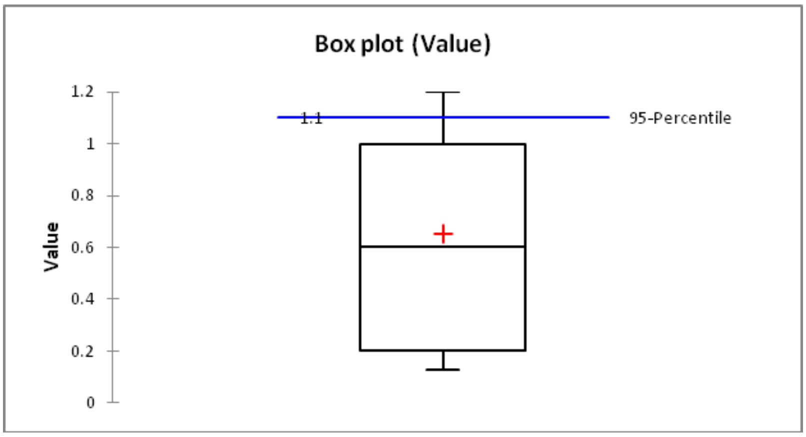


Figure 25 – Box plot for Cd Urban

Table 6: 95th percentile values and their confidence limits for Cd Urban, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	**P H**	Robust	R L	R H
48	1.1	1.1	1.2	1.1	1.00	**1.20**	0.70	0.69	0.72

Parametric percentiles were fitted to this data set as it showed a symmetrical normal distribution. See Appendix 7 for further information.

The NBC for the cadmium urban domain is 1.2mg/kg.

Copper (Cu)

Copper is a naturally occurring metal, which is found in plants and animals as well as in the ground, the water and the air, and is essential for life in low concentrations. Natural sources of copper include volcanoes, windblown dust, decaying vegetation and forest fires. Anthropogenic sources of copper are typically from agricultural practices (GAC report, Nathanail *et al*, 2009).

The NBC of copper was not expected to be high in the Borough of Darlington given the natural and anthropogenic sources, the results below agree with this.

Rural Domain (Cu)

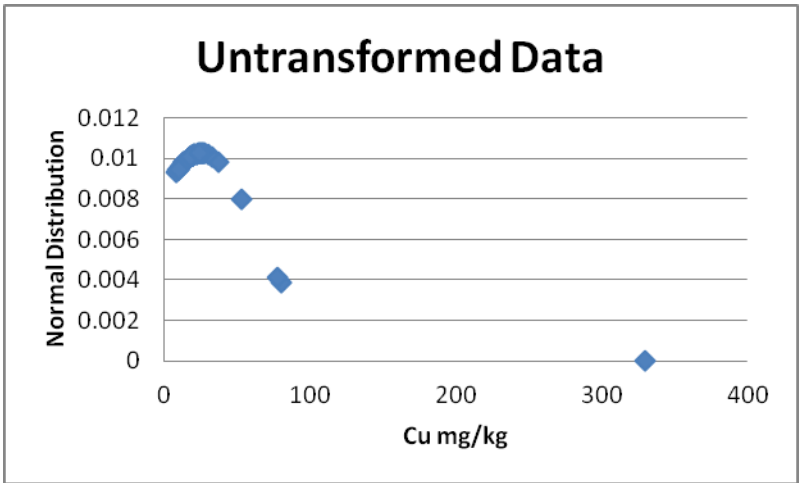


Figure 26 – Normal distribution plot for the untransformed data set for Cu rural, showing a skewed distribution

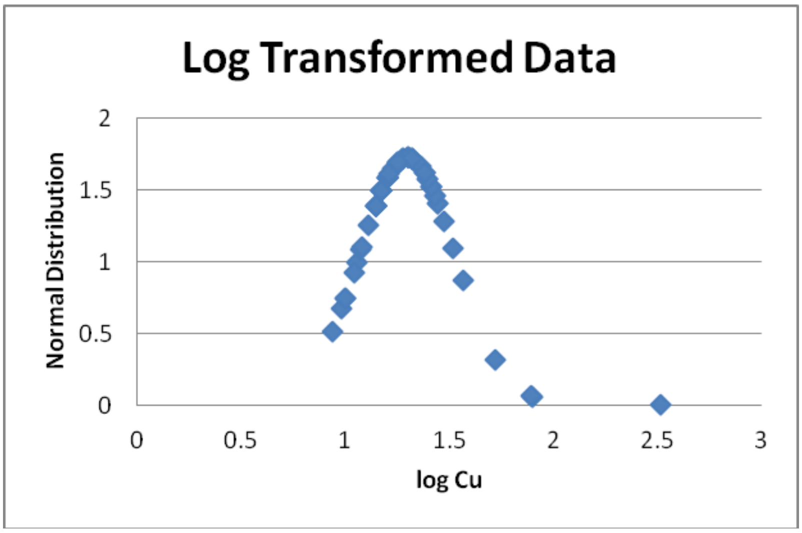


Figure 27 – Normal distribution plot for the log transform of the Cu Rural data set, showing a Gaussian distribution with outliers in the right hand tail

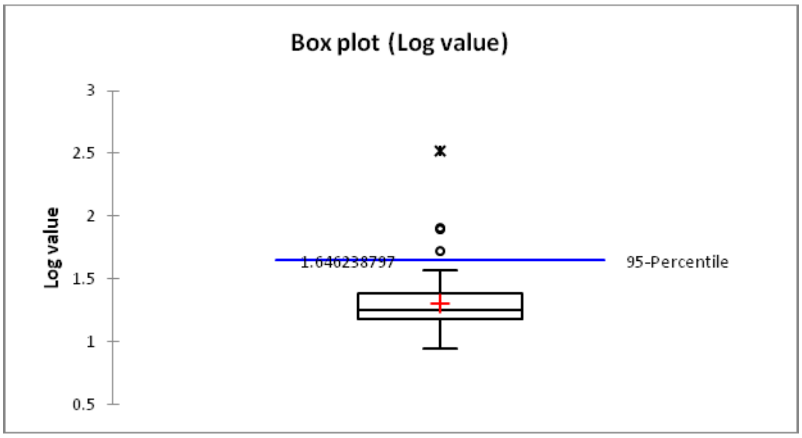


Figure 28 – Box plot of the log transformed data for Cu Rural

Table 7: 95th percentile values and their confidence limits for Cu Rural, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	**R H**
70	44.28	28	330	45.09	39.80	51.07	21.40	21.04	**21.77**

Robust percentiles have been fitted to the log transformation of this data set because it displayed a Gaussian distribution but with outliers in the right hand tail. Once the percentiles had been fitted to the transformed data inverse logs were applied to all the values in order to back transform and get the value of the NBC. See Appendix 8 for further information.

The NBC for the copper rural domain is 21.8mg/kg.

Urban Domain (Cu)

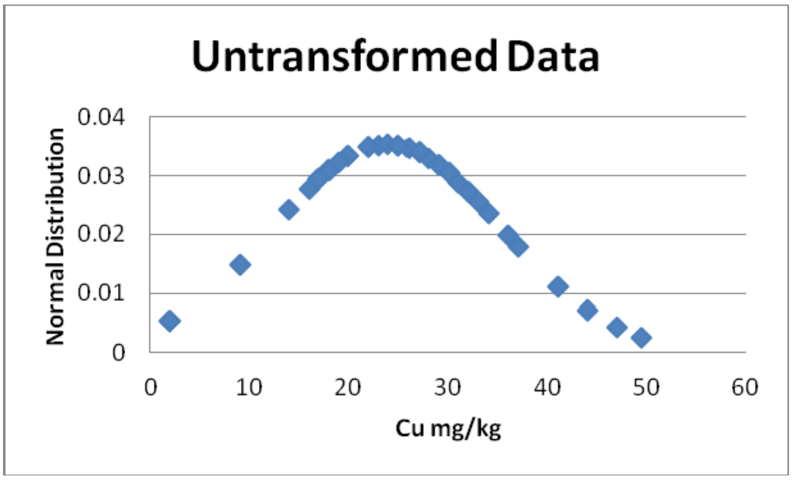


Figure 29 – Normal distribution plot for the untransformed data set for Cu Urban, showing a symmetrical distribution

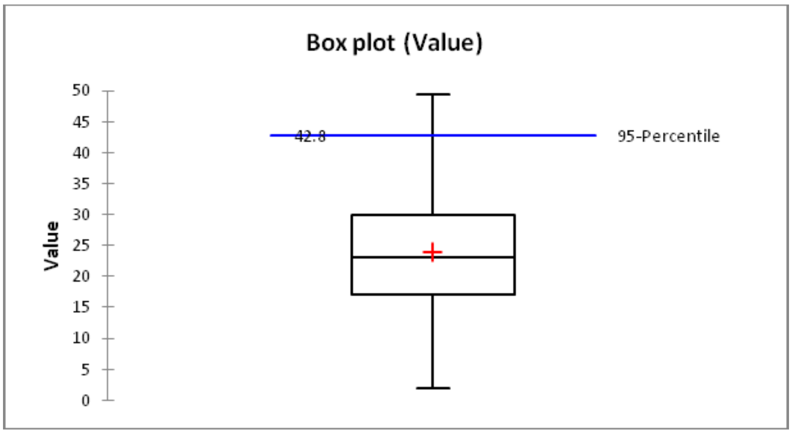


Figure 30 – Box plot for Cu Urban

Table 8: 95th percentile values and their confidence limits for Cu Urban, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	**P H**	Robust	R L	R H
48	42.8	37	49.5	42.95	39.75	**46.15**	24.19	24.09	24.29

Parametric percentiles have been fitted to this data set as it displayed a symmetrical normal distribution plot. See Appendix 9 for further information.

The NBC for the copper urban domain is 46.2mg/kg.

Mercury (Hg)

Mercury is a heavy metal, which can be found in the environment due to natural as well as anthropogenic processes. Natural inputs are associated with volcanic activity and microbial breakdown of inorganic mercury into mercury compounds, which are found in soils. Anthropogenic sources of mercury are through combustion of fossil fuels and fertilisers (The Environment Agency, 2009).

The NBC of arsenic was not expected to be high in the Borough of Darlington given the natural and anthropogenic sources, the results below agree with this.

Rural Domain (Hg)

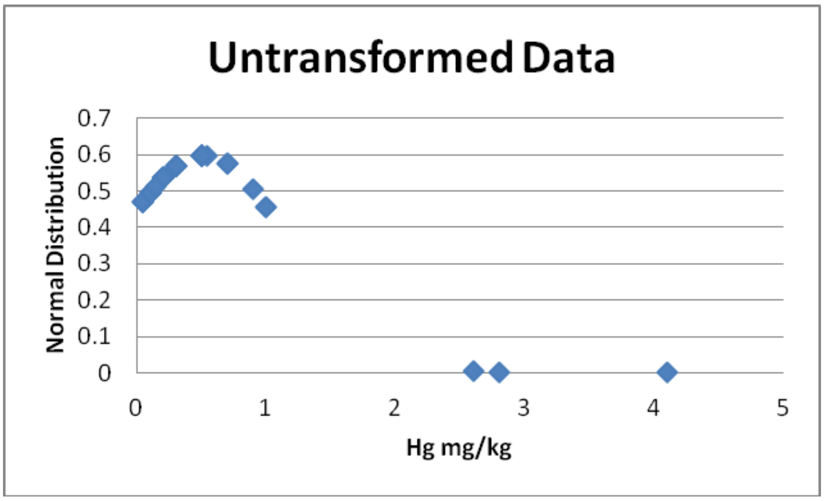


Figure 31 – Normal distribution plot for the untransformed data set for Hg Rural, showing a Gaussian distribution with outliers in the right hand tail

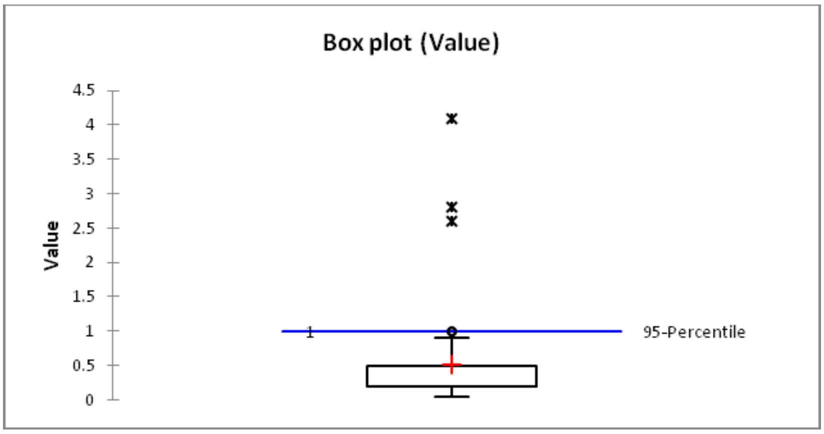


Figure 32 – Box plot for Hg Rural

Table 9: 95th percentile values and their confidence limits for Hg Rural, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	**R H**

61	1	0.7	4.1	1	0.83	1.17	0.52	0.52	**0.52**

Robust percentiles give the value of the NBC for this data set because it showed a Gaussian distribution with outliers in the right hand tail. See Appendix 10 for further information.

The NBC for the mercury rural domain is 0.52mg/kg.

Urban Domain (Hg)

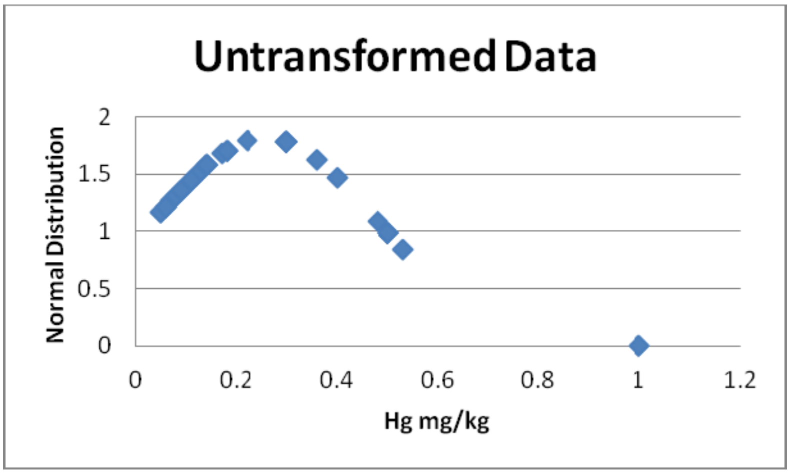


Figure 33 – Normal distribution plot for the untransformed data set for Hg Urban, showing a skewed distribution

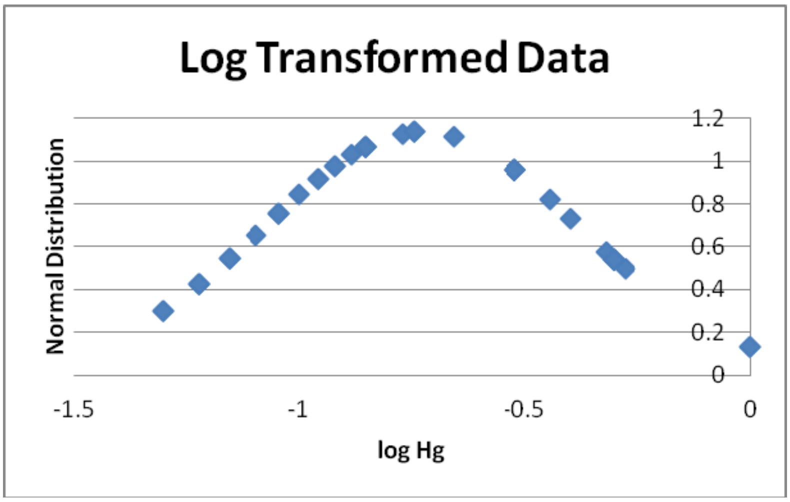


Figure 34 – Normal distribution plot for the log transformed data for Hg Urban, showing a symmetrical distribution

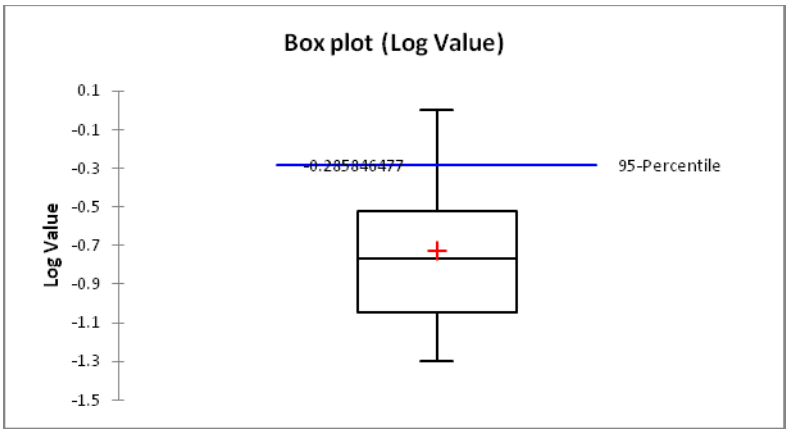


Figure 35 – Box plot for the log transformed data for Hg Urban

Table 10: 95th percentile values and their confidence limits for Hg Urban, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	**P H**	Robust	R L	R H
48	0.52	0.5	1	0.52	0.41	**0.65**	0.20	0.20	0.20

Parametric percentiles were fitted to the log transform of this data set as it showed a symmetrical distribution. Once the percentiles had been fitted inverse logs were applied to the values in order to get the actual value of the NBC. See Appendix 11 for further information.

The NBC for the mercury urban domain is 0.65mg/kg.

Nickel (Ni)

Nickel is a metal, which occurs naturally in soil but also has a number of anthropogenic inputs. Nickel occurs naturally in soils due to the weathering of the underlying geology. The rocks with the highest concentration of nickel are basic igneous rocks; sedimentary rocks have much lower levels. The soil forming process can also have a big effect on the amount of nickel in the soil, clays, silts and fine-grained loams have higher concentrations than coarse-grained loams, sandy and peaty soils.

The underlying geology of the Borough of Darlington is all sedimentary and so does not lend itself to having high levels of nickel either in the urban or rural domain. Darlington soils are predominantly clay, with occasional sandy soils, which are more frequent in the rural area. This soil type could potentially lend itself to elevated levels of nickel, but there will not be a high level of nickel from the underlying sedimentary geology transferred to the soil in the Borough of Darlington. Nickel in soils is widespread due to deposition from the burning of fossil fuels and fertilisers (The Environment Agency, 2009).

The NBC of nickel was not expected to be high in the Borough of Darlington given the natural and anthropogenic sources, the results below agree with this. The results below show that the concentration of nickel is higher in rural areas of Darlington than in urban.

Rural Domain (Ni)

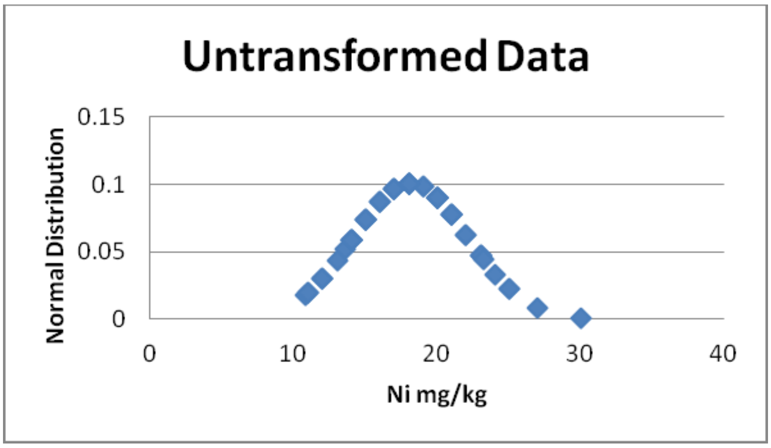


Figure 36 – Normal distribution plot for the untransformed data for Ni Rural, showing a symmetrical distribution

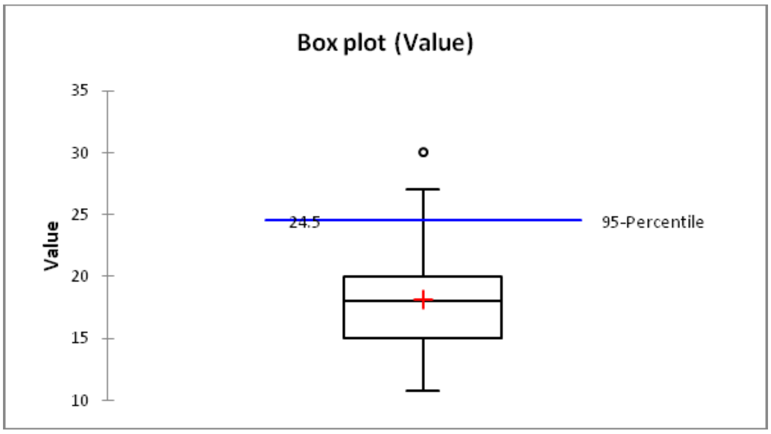


Figure 37 – Box plot for Ni Rural

Table 11: 95th percentile values and their confidence limits for Ni Rural, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	**P H**	Robust	R L	R H
70	24.5	23	30	24.55	23.62	**25.48**	18.25	18.22	18.27

Parametric percentiles were used to calculate the NBC of this data set because it showed a symmetrical distribution. See Appendix 12 for further information.

The NBC for the nickel rural domain is 25.5mg/kg.

Urban Domain (Ni)

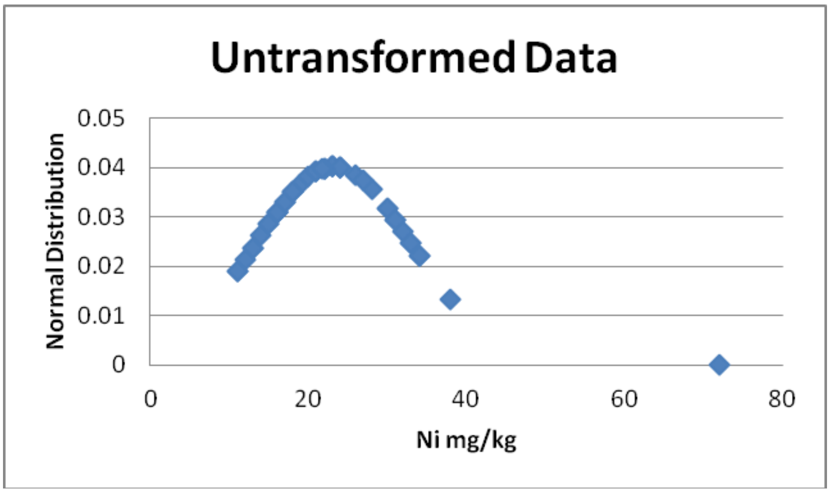


Figure 38 – Normal distribution plot for the untransformed data for Ni Urban, showing a Gaussian distribution with outliers in the right hand tail

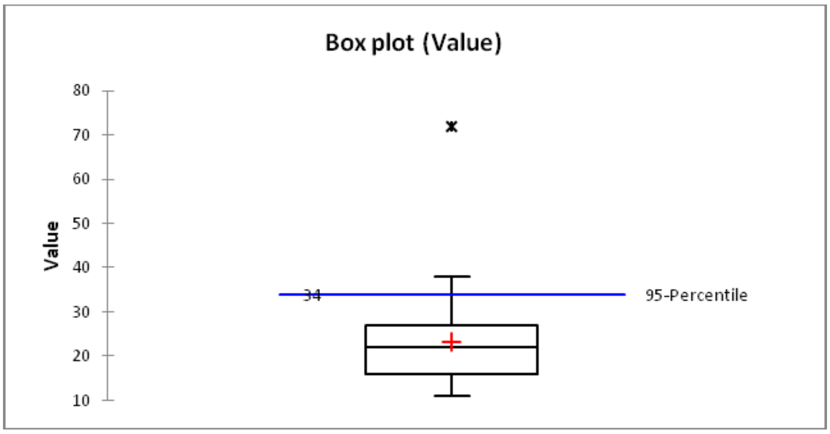


Figure 39 – Box plot for Ni Urban

Table 12: 95th percentile values and their confidence limits for Ni Urban, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	**R H**
48	34	33	72	34	31.19	36.81	24.37	24.04	**24.71**

Robust percentiles were used to fit the NBC to this data set as it displayed a Gaussian distribution with outliers in the right hand tail. This method of fitting percentiles is more conservative than fitting parametric percentiles due to the nature of the spread of the data set; this could be why this project has found that the rural values for nickel are higher than the urban values. See Appendix 13 for further information.

The NBC for the nickel urban domain is 24.7mg/kg.

Lead (Pb)

Lead is a naturally occurring component that is found in rocks. It is derived from igneous rocks and also found in high levels in sedimentary rocks such as shale and mudstone from the redistribution of weathered igneous sediment. Therefore Darlington's underlying geology does not lend itself to producing high lead values (The Environment Agency, 2002).

Anthropogenic sources of lead include sewage sludge used as fertiliser and vehicle exhaust fumes. Due to the volume of traffic in the town centre it was considered likely that the urban area would have higher levels of lead than the rural area. The results below confirm this prediction.

Rural Domain (Pb)

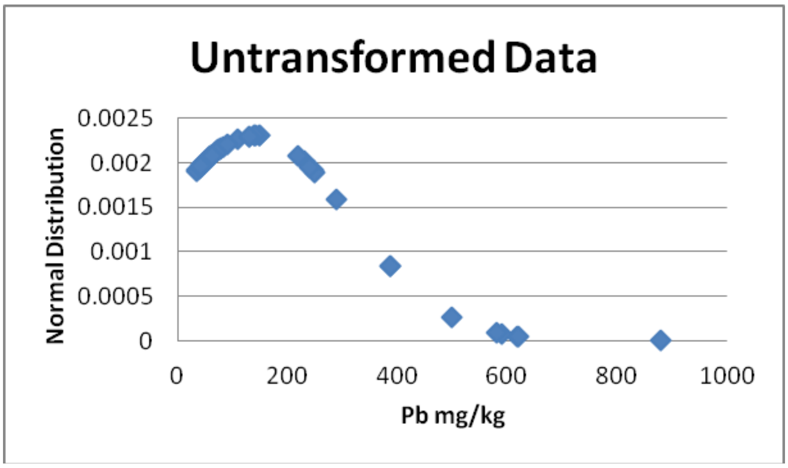


Figure 40 – Normal distribution plot for the untransformed data set for Pb Rural, showing a skewed distribution

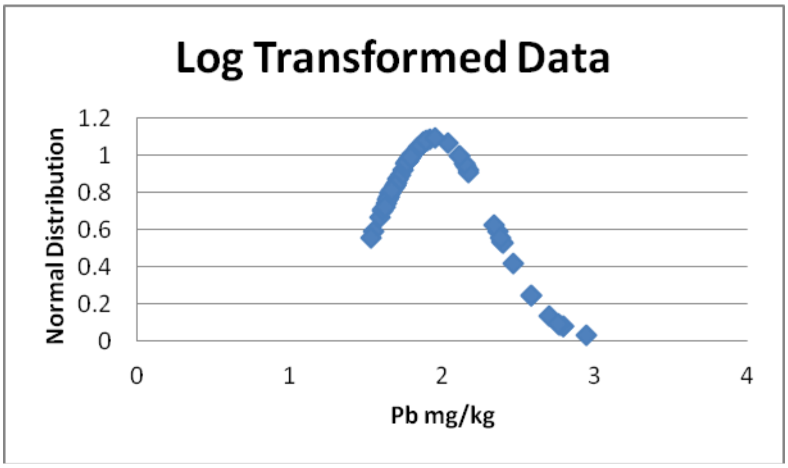


Figure 41 – Normal distribution plot for the log transformation of the data set for Pb Rural, still showing a skewed distribution

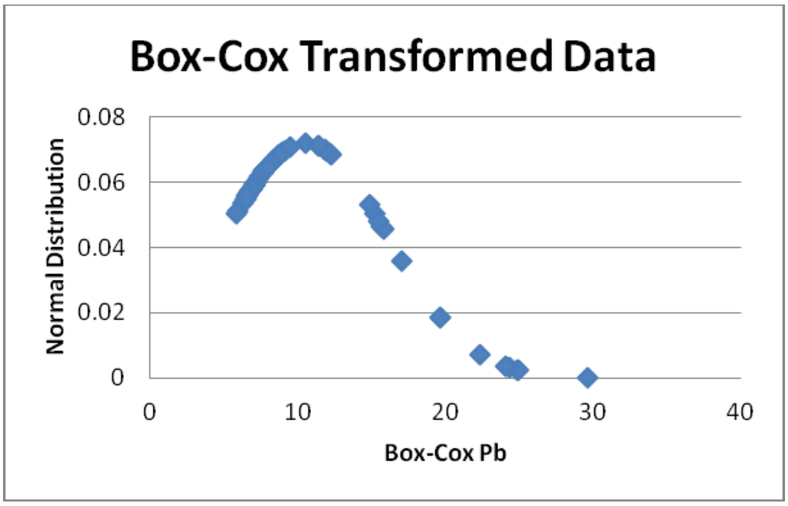


Figure 42 – Normal distribution plot for the Box-Cox transformed data set for Pb Rural, still showing a skewed distribution

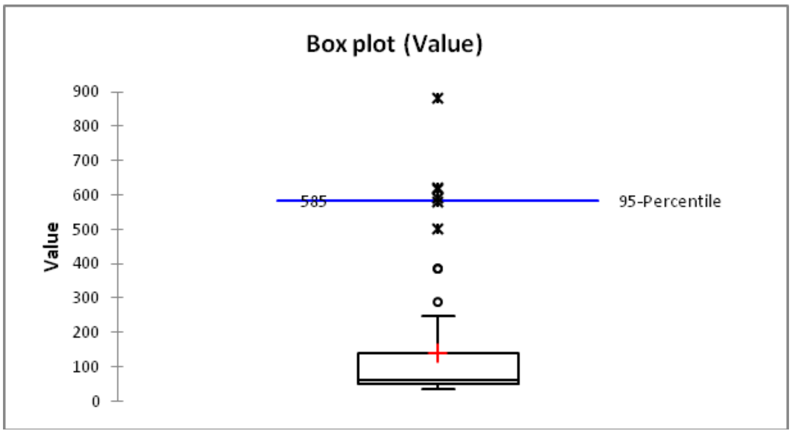


Figure 43 – Box plot for the original data for Pb Rural

Table 13: 95th percentile values and their confidence limits for Pb Rural, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	**R H**
70	585	290	880	585.5	545.0	626.1	217.9	199.8	**236.1**

Following the BGS guidelines empirical percentiles should have been fitted to this data set since it did not approximate a symmetrical or Gaussian normal distribution curve after any of the transformations. However, the upper 95% confidence limit of the upper 95th percentile is given as the maximum value of the data set; this was not deemed an appropriate value to assign to the NBC after looking at the normal distribution plot.

Since this data set was heavily influenced by outliers in the right hand tail it was decided, after consideration, to instead select the value as given by the robust percentile method. The robust percentile method is robust to outliers so gave a more reasonable value for the NBC. See Appendix 14 for further information.

The NBC for the lead rural domain is 236mg/kg.

Urban Domain (Pb)

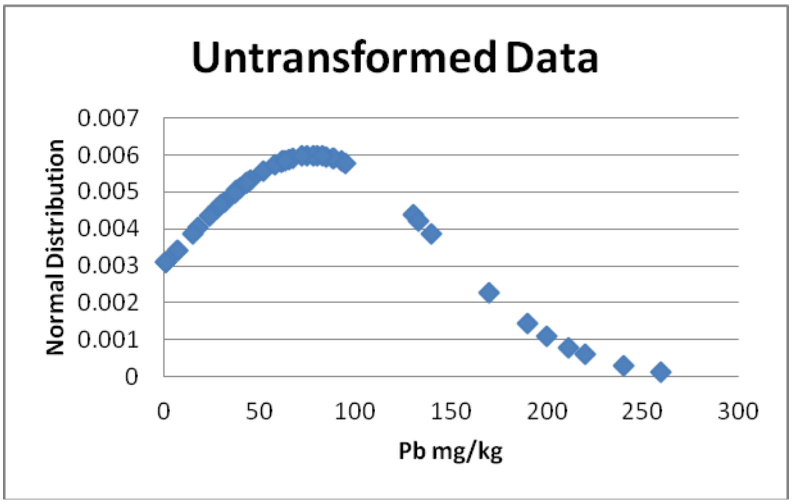


Figure 44 – Normal distribution plot for the untransformed data set for Pb Urban, showing a skewed distribution

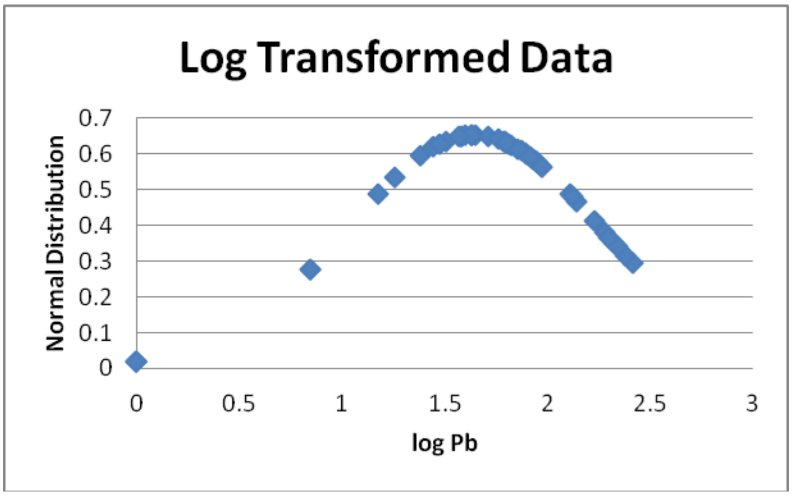


Figure 45 – Normal distribution plot for the log transformed data set for Pb Urban, showing a symmetrical distribution

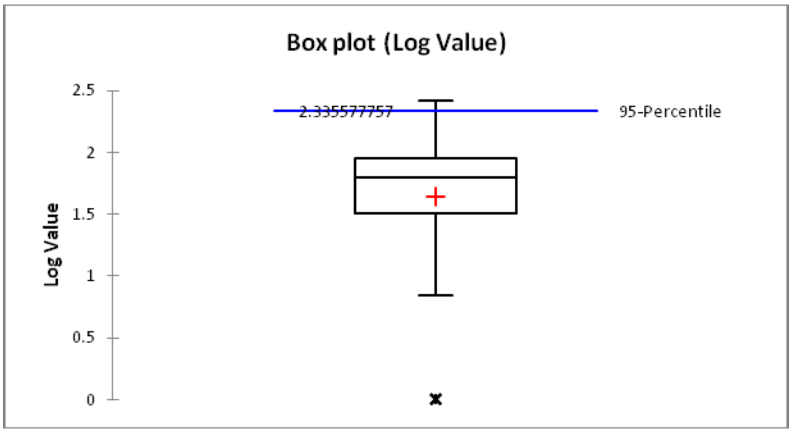


Figure 46 – Box plot for the log transformed data set for Pb Urban

Table 14: 95th percentile values and their confidence limits for Pb Urban, where L and U are lower and upper percentiles respectively

No. of samples	Empirical	Emp L	Emp H	Parametric	P L	**P H**	Robust	R L	R H
48	216.6	190	260	217.0	145.7	**323.1**	122.0	110.9	134.3

Parametric percentiles have been fitted to the log transform of the data set as it displayed a symmetrical normal distribution. Once the percentiles had been fitted to the log transformed data inverse logs were applied to all of the values in order to get the actual value of the NBC. See Appendix 15 for further information.

The NBC for the lead urban domain is 323mg/kg.

Limitations of Project

This project has a number of limitations associated with it. Where possible the impact of these limitations has been reduced as much as possible, but the nature of the analysis means that they cannot simply be ignored and may have had an effect on the calculated NBC values.

The data itself has limitations because samples with low values of certain contaminants cannot be given an exact value due to being below the detection limit. This has occurred for all of the contaminants, except for nickel. In these cases the value was displayed as <*value* in the site investigation. For the purposes of this project the limit of detection was used as the less than value, for example BaP < 0.1 = 0.1, otherwise it could not be processed in the statistical analysis. In the case of benzo[a]pyrene in the rural domain this limitation led to only half of a bell curve in the normal distribution so the data set could not be approximated as Gaussian even after transformations.

The criteria for determining sites suitable for inclusion in the project, topsoil samples of greenfield sites, eliminated a number of site investigations held by the Council, which were looked at with respect to this project. Once the data had been condensed into useable samples the locations of the sampling points were plotted on an ArcGIS layer in order to show their spatial distribution across the Borough. Sample location maps in reports and grid references from sample logs were used to plot the sample points on the ArcGIS layer. During this process it was found that some grid references were wrong. In these cases the sample location map, which shows the position of samples in the area of a site investigation was used and although the samples were still included, the grid references were not included in the spreadsheet. For sample points taken from a site-derived stockpile of topsoil all of the sample points were plotted in the same location. This will not have had a major effect with respect to showing the general spread of the data but it should be noted that not all of the sample points will be in precisely the right location as it was done by eye.

Any limitations in the statistical analysis will come from the programs chosen to perform the statistics in, as well as personal judgement of the results. For example, when calculating the robust percentiles for log values, many times the median±2MAD returned negative values, or the median-2MAD value was higher than the median+2MAD value due to a negative MAD. The higher of the two values was always used as the 95^{th} percentile, as logically it should be the highest value. There was no guidance in the Defra/BGS (2012) report for what to do in case of these kinds of values so this method was deemed to be the most sensible approach for calculating robust percentiles.

The decision to use robust percentiles instead of empirical percentiles for the rural domains of benzo[a]pyrene and lead came as a result of perceived limitations in the method of working out the empirical percentiles. Using the XLSTAT programme the upper confidence limit for the 95th percentile often came out at the maximum value in the data set. It was assumed this came from the fact that our data sets were limited in numbers, even though they exceeded the required value for this project as set by the Defra, BGS (2012) report. The empirical results produced from XLSTAT took all the outliers into account which was deemed inappropriate to use for a distribution, which could not approximate as Gaussian.

The limitations of this project were counteracted by careful consideration of the data set, and by only including greenfield samples; therefore Darlington Borough Council has confidence in the NBC's calculated.

Conclusion

Table 15: Comparing this project with BGS data and GAC/SGV levels

Contaminant	This project (rural) (mg/kg)	This project (urban) (mg/kg)	BGS rural (mg/kg)	BGS urban (mg/kg)	GAC/SGV residential (mg/kg)
Arsenic	8.3	14.1	32	*	32 (SGV)
Benzo[a]pyrene	0.25	3.2	0.5	3.6	0.83 (GAC)
Cadmium	0.54	1.2	1.0	2.1	10 (GAC)
Copper	21.8	46.2	62	190	2330 (GAC)
Mercury (elemental)	0.52	0.65	0.5	1.9	1 (SGV)
Nickel	25.5	24.7	42	**	130 (SGV)
Lead	236	323	180	820	450 (SGV)

* No value was given for the urban domain of arsenic in the BGS report, the mineralisation domain, with a value of 290mg/kg, and the ironstone domain, with a value of 220mg/kg, were given instead.
** No value was given for the urban domain of nickel in the BGS report, the ironstone domain, with a value of 230, and the Peak District domain, with a value of 120mg/kg, were given instead.

The NBCs for arsenic, cadmium, copper, mercury, nickel and lead calculated in this project are much more conservative than those calculated by the BGS and the published SGV/GACs. It is therefore likely that the current SGV/GAC values will continue to be used as the screening levels, which denotes whether further risk assessment is required.

The only contaminant, which the NBC exceeds the GAC is benzo[a]pyrene. The GAC for benzo[a]pyrene is 0.83mg/kg for residential 1% Soil Organic Matter (GAC Report, Nathanail *et al.*, 2009). In line with the Part 2A Contaminated Land Statutory Guidance, NBCs of contaminants in soil should not be considered to cause land to qualify as contaminated land (pose an unacceptable risk), unless there is a particular reason to consider otherwise. Under the Contaminated Land Inspection Strategy for Borough of Darlington it is stated that: if necessary, normal background concentrations will be used as a guide as to what are reasonable levels to support the decision of whether land within the Borough is contaminated land under Part 2A. Darlington Borough Council is confident that the NBCs calculated are typical for the Borough of Darlington and therefore there is no reason to consider the NBC of 3.2mg/kg for urban benzo[a]pyrene to pose an unacceptable risk.

These conclusions are however subject to the release of Category 4 screening levels.

References

Brock S.E., n.d. Descriptive statistics and psychological testing. [online] CSUS. available at www.csus.edu/.../Descrptive%20Statistics%20and%20the%20Normal%20Curve.pdf [accessed 23 August 2013]

Cave, MR; Johnson, CC; Ander, EL and Palumbo-Roe, B. 2012. Methodology for the determination of normal background contaminant concentrations in English soils. *British Geological Survey Commissioned Report*, British Geological Survey, Nottingham

Darlington Borough Council. 2013. Contaminated Land Inspection Strategy, The Borough of Darlington, Part 2A of the Environmental Protection Act 1990. Darlington Borough Council. Darlington.

Defra, 2012. Environmental Protection Act 1990: Part 2A Contaminated Land Statutory Guidance. April 2012. Department for Environment, Food and Rural Affairs (Defra). HM Government.

Johnson, CC; Ander, EL; Cave, MR and Palumbo-Roe, B. 2012. Normal background concentrations (NBCs) of contaminants in English soils: Final project report. *British Geological Survey Commissioned Report*, British Geological Survey, Nottingham

Knowware, 2013. Box Cox Transformation in Excel. [online] available at http://www.qimacros.com/quality-tools/boxcox/ [accessed 23 August 2013]

Nathanail, C.P., McCaffrey, C., Ashmore, M.H., Cheng, Y.Y., Gillett, A., Ogden, R. and Scott, D. 2009. The LQM/CIEH Generic Assessment Criteria for Human Health Risk Assessment (2nd edition). Land Quality Press, Nottingham.

The Environment Agency, 2002. Soil Guideline Values for Lead Contamination. The Environment Agency, Bristol.

The Environment Agency, March 2009. Soil Guideline Values for mercury in soil. The Environment Agency, Bristol.

The Environment Agency, March 2009. Soil Guideline Values for nickel in soil. The Environment Agency, Bristol.

The Environment Agency, May 2009. Soil Guideline Values for inorganic arsenic in soil. The Environment Agency, Bristol.

The Higher Education Academy, 2009. Probability and statistics in Microsoft Excel. [online] available at http://msor.rsscse.org.uk/leaflets/psme/psme.pdf [accessed 23 August 2013]

XLSTAT, 2013. Calculating quantiles or percentiles with XLSTAT. [online] available at http://www.xlstat.com/en/learning-center/tutorials/quantiles-or-percentiles-with-xlstat.html [accessed 23 August 2013]

Appendices

Appendix 1 – Detailed Methodology

This is a detailed description of the statistical analysis carried out on the data sets.

The initial stage of the statistical analysis was to produce normal distributions of the data sets. This was done in excel using the NORM.DIST(x, mean, standard_dev, cumulative) function, where x was a value in the data set. This was worked out for each data point in the set and then the NORM.DIST value was plotted against the original value in a scatter graph.

The next stage of the analysis was to calculate the SC and OS using the following equations:

$$SC = \sum(xi-\mu)^3/N\sigma^3$$

Where μ is the mean, σ is the standard deviation and N is the number of data points

$$OS = ((Q_{0.875}-Q_{0.5})-(Q_{0.5}-Q_{0.125}))/(Q_{0.875}-Q_{0.125})$$

Where Qn is the nth quantile of the data set

These calculations were also done in excel. The SC calculation only required the mean and standard deviation of the data set to be calculated. This was done using the =AVERAGE and =STDEV.S functions in excel, =STDEV.S was used as it was defined as giving the standard deviation for a sample. The equation $=(xi-\mu)^3/N\sigma^3$ was applied to each value in the data set and then the sum was calculated to get the SC value.

The OS calculation required more steps, the 87.5th percentile, the 12.5th percentile and the median needed calculating. This was done using the =QUARTILE and =PERCENTILE functions in excel, once these had been calculated the OS calculation was computed. This way of producing percentiles is the parametric method so the parametric 95th percentile of the data was also calculated at the same time.

The OS and SC values were used to classify data sets as detailed in the main report.

Where a log transformation was required the function =LOG10 in excel was applied to each value within the data set.

Where a Box-Cox transformation was required a basic excel package could not produce this equation. The add-in QI Macros for excel was used to produce this transformation. After selecting the data, the Tools button was selected, then the Data Transformation button, then the Box-Cox Transformation button (Knowware, 2013). The excel formula produced by the add-in came out as

=IF(ISNUMBER(x),IF(ƛ>0,(x^ƛ),IF(ƛ<0,1/(x^-ƛ),LN(x))),x) where ƛ = 0.5.

After the transformations the three different types of percentiles had to be fitted to the data. Parametric percentiles were fitted using the =PERCENTILE(array, k) function in excel.

To get robust percentiles the BGS methodology states that “percentiles are fitted using median and the median absolute deviation (MAD) in place of the mean and standard deviation as these measures are robust to outliers”. Using the mean and standard deviation approximately 95% of the data set lies within 2 standard deviations of the mean (Brock, n.d.), by fitting percentiles using the median and 2MAD instead, a robust value for the 95th percentile was calculated. Robust percentiles were therefore calculated by finding the median and MAD, these were both found in excel. The median is found using the =QUARTILE function and the MAD is found by computing (∑median-x)/number of values. Median+2MAD and Median-2MAD were then calculated and the highest value was taken as the 95th percentile.

The empirical percentiles could not be calculated using a basic excel sheet, so the add-in XLSTAT was used. Clicking the data transformation button and selecting “quantiles estimation” allowed the calculation of the empirical percentiles. The charts tab in the wizard allowed the selection of all boxes in order to display the graphs and typing 95 in the box labelled “Show quantile on charts (%)” labelled the 95th percentile on the charts (XLSTAT 2013). This produced a new page with all the empirical percentiles detailed on it, including charts. Using this add-in the upper 95% confidence limits were also displayed for the empirical values.

In order to work out the upper and lower confidence limits for parametric and robust percentiles basic excel functions were used. For parametric percentiles the formula used was: =x±CONFIDENCE.NORM(alpha, standard_dev,size), where x was the value of the 95th percentile and alpha = 1-confidence limit, so 1-0.95 = 0.05. For robust percentiles the equation was the same in excel, but the value of the standard deviation was replaced with the MAD value. CONFIDENCE.NORM was used because the distributions were approximating symmetrical normal distributions.

The excel equation for confidence limits was determined by modifying an equation for the confidence limits on the mean given as =AVERAGE(x)±CONFIDENCE(alpha, sigma, COUNT(x)) (The Higher Education Academy, 2009). By replacing the average value with that of the 95th percentile it was possible able to work out the confidence limits on the 95th percentile instead of the mean.

Appendix 2 – Arsenic Rural Data

Full spreadsheet for arsenic rural, including parametric percentiles calculation

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
18	Arsenic	R	1	12	61	7.886885246	2.14230992	4.3	5.9	6.9	8	8.7	9.3	12	16	0.116020771	1.541453208	1.620267624	-0.23529412	Gaussian	11.46239	12.53761
19	Arsenic	R	2	8.3												0.000117553						
20	Arsenic	R	3	5.7												-0.017438184						
21	Arsenic	R	4	8.3												0.000117553						
22	Arsenic	R	5	8.7												0.000896352						
23	Arsenic	R	6	8.1												1.61385E-05						
32	Arsenic	R	7	4.3												-0.076944169						
33	Arsenic	R	8	4.7												-0.053966288						
35	Arsenic	R	9	7												-0.001163123						
36	Arsenic	R	10	6												-0.011201076						
37	Arsenic	R	11	7												-0.001163123						
38	Arsenic	R	12	8												2.41313E-06						
39	Arsenic	R	13	8												2.41313E-06						
40	Arsenic	R	14	9												0.002299551						
41	Arsenic	R	15	8												2.41313E-06						
42	Arsenic	R	16	5.8												-0.01515371						
43	Arsenic	R	17	6.9												-0.001602593						
44	Arsenic	R	18	6.1												-0.009512907						
45	Arsenic	R	19	6												-0.011201076						
46	Arsenic	R	20	6.5												-0.004447796						
64	Arsenic	R	21	6												-0.011201076						
65	Arsenic	R	22	7												-0.001163123						
66	Arsenic	R	23	8												2.41313E-06						
67	Arsenic	R	24	8												2.41313E-06						
68	Arsenic	R	25	8												2.41313E-06						
69	Arsenic	R	26	8												2.41313E-06						
70	Arsenic	R	27	8												2.41313E-06						
71	Arsenic	R	28	9												0.002299551						
72	Arsenic	R	29	8												2.41313E-06						
73	Arsenic	R	30	9												0.002299551						
74	Arsenic	R	31	10												0.015732311						
75	Arsenic	R	32	7												-0.001163123						
76	Arsenic	R	33	10												0.015732311						
77	Arsenic	R	34	8												2.41313E-06						
78	Arsenic	R	35	9												0.002299551						
79	Arsenic	R	36	9												0.002299551						
80	Arsenic	R	37	8												2.41313E-06						
81	Arsenic	R	38	8												2.41313E-06						
82	Arsenic	R	39	7												-0.001163123						
83	Arsenic	R	40	8												2.41313E-06						
84	Arsenic	R	41	13												0.22288451						
85	Arsenic	R	42	6.8												-0.002140791						
86	Arsenic	R	43	7.4												-0.000192443						
87	Arsenic	R	44	7.3												-0.000337041						
88	Arsenic	R	45	7.3												-0.000337041						
89	Arsenic	R	46	7.6												-3.93684E-05						
93	Arsenic	R	47	6.1												-0.009512907						
94	Arsenic	R	48	5.5												-0.022673465						
97	Arsenic	R	49	5												-0.040115465						
98	Arsenic	R	50	5.4												-0.02564428						
103	Arsenic	R	51	7.6												-3.93684E-05						
104	Arsenic	R	52	9.6												0.008382675						
106	Arsenic	R	53	10.1												0.018073201						
111	Arsenic	R	54	9												0.002299551						
112	Arsenic	R	55	15												0.600071102						
113	Arsenic	R	56	7												-0.001163123						
114	Arsenic	R	57	5												-0.040115465						
115	Arsenic	R	58	16												0.890401993						
116	Arsenic	R	59	7												-0.001163123						
117	Arsenic	R	60	7												-0.001163123						
118	Arsenic	R	61	9												0.002299551						

Arsenic rural robust percentile calculation spreadsheet

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
61	12	8	-4	6.9	0.113115	8.226229508	7.773770492	8.197844	8.254615
	8.3		-0.3						
	5.7		2.3						
	8.3		-0.3						
	8.7		-0.7						
	8.1		-0.1						
	4.3		3.7						
	4.7		3.3						
	7		1						
	6		2						
	7		1						
	8		0						
	8		0						
	9		-1						
	8		0						
	5.8		2.2						
	6.9		1.1						
	6.1		1.9						
	6		2						
	6.5		1.5						
	6		2						
	7		1						
	8		0						
	8		0						
	8		0						
	8		0						
	8		0						
	9		-1						
	8		0						
	9		-1						
	10		-2						
	7		1						
	10		-2						
	8		0						
	9		-1						
	9		-1						
	8		0						
	8		0						
	7		1						
	8		0						
	13		-5						
	6.8		1.2						
	7.4		0.6						
	7.3		0.7						
	7.3		0.7						
	7.6		0.4						
	6.1		1.9						
	5.5		2.5						
	5		3						
	5.4		2.6						
	7.6		0.4						
	9.6		-1.6						
	10.1		-2.1						
	9		-1						
	15		-7						
	7		1						
	5		3						
	16		-8						
	7		1						
	7		1						
	9		-1						

Arsenic rural empirical calculation spreadsheet

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 15:52:09
Data: Workbook = As Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$62 / 61 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

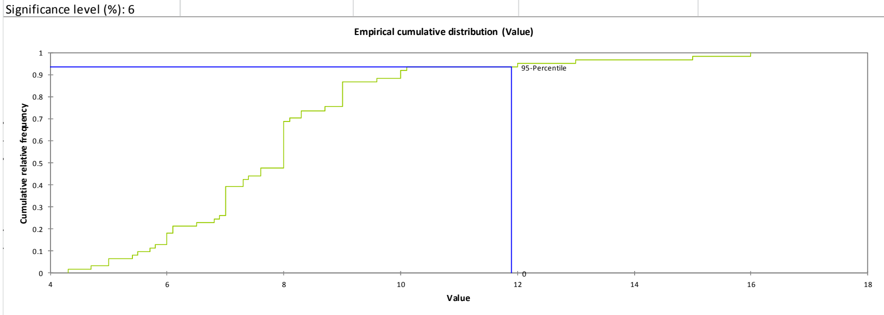
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	61	0	61	4.300	16.000	7.887	2.142

Percentile table (Weighted average at x(Np)):

Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	16.000				
99%	15.390	13.000	16.000	13.000	16.000
95%	11.905	10.000	16.000	9.000	16.000
90%	9.960	9.000	13.000	9.000	15.000
3rd Quartile 75%	8.600	8.000	9.000	8.000	9.000
Median 50%	8.000	7.000	8.000	7.000	8.000
1st Quartile 25%	6.825	6.000	7.000	5.800	7.000
10%	5.520	4.700	6.000	4.300	6.000
5%	5.000	4.300	5.500	4.300	5.400
1%	4.544	4.300	4.700	4.300	4.700
Minimum 0%	4.300				

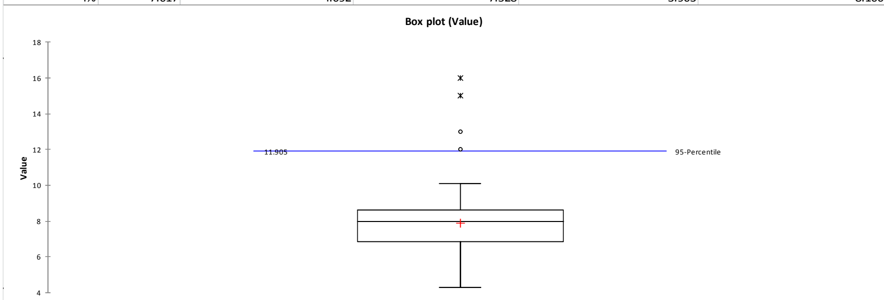
Value of the 95-percentile: 11.905

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 15:52:10
Data: Workbook = As Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$62 / 61 rows and 1 column
Significance level (%): 6



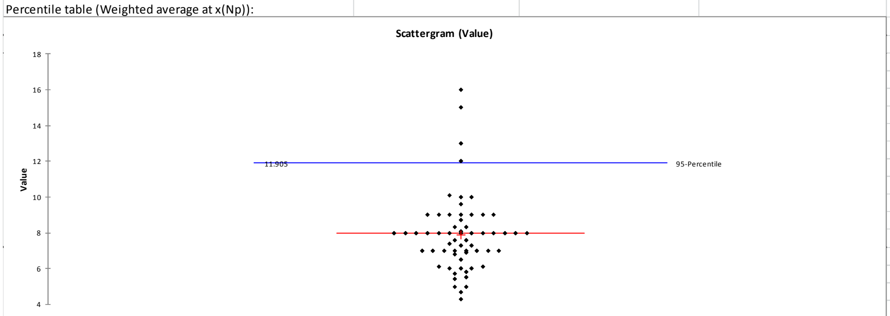
Mean	Std. deviation
7.887	2.142

4th Quartile 75%	-3.383	-0.967	-2.893	-0.878	-3.256
Median 50%	-4.594	-2.008	-4.438	-1.886	-4.872
2nd Quartile 25%	-5.806	-3.050	-5.983	-2.894	-6.489
-4%	-7.017	-4.092	-7.528	-3.903	-8.106



						Mean	Std. deviation
Value	61	0	61	4.300	16.000	7.887	2.142

Percentile table (Weighted average at x(Np)):



XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 15:52:12

Appendix 3 – Arsenic Urban Data

Full spreadsheet for arsenic urban, including parametric percentiles calculation

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev.	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
1	Arsenic	U	1	11.9	48	8.165833	3.998993	2	2	6.15	8	11	12	13	23	0.016962419	0.752402811	0.801820572	-0.2	Symmetrical	11.8687	14.1313
2	Arsenic	U	2	7												-0.000516198						
3	Arsenic	U	3	7												-0.000516198						
4	Arsenic	U	4	7												-0.000516198						
5	Arsenic	U	5	8												-1.48567E-06						
6	Arsenic	U	6	8												-1.48567E-06						
7	Arsenic	U	7	6.2												-0.002474838						
8	Arsenic	U	8	5.9												-0.003789576						
9	Arsenic	U	9	6.8												-0.000830042						
10	Arsenic	U	10	6.3												-0.002116047						
11	Arsenic	U	11	6.9												-0.000660749						
12	Arsenic	U	12	6												-0.003309648						
13	Arsenic	U	13	7.9												-6.11977E-06						
14	Arsenic	U	14	6.3												-0.002116047						
15	Arsenic	U	15	14												0.0646909						
16	Arsenic	U	16	23												1.063398791						
17	Arsenic	U	17	9												0.000189089						
24	Arsenic	U	18	8												-1.48567E-06						
25	Arsenic	U	19	8												-1.48567E-06						
26	Arsenic	U	20	10												0.002010125						
27	Arsenic	U	21	9												0.000189089						
28	Arsenic	U	22	7												-0.000516198						
29	Arsenic	U	23	5.2												-0.008498591						
30	Arsenic	U	24	6.5												-0.001505918						
31	Arsenic	U	25	4.5												-0.016048145						
34	Arsenic	U	26	6												-0.003309648						
47	Arsenic	U	27	11.8												0.01563584						
48	Arsenic	U	28	10.9												0.00665858						
50	Arsenic	U	29	2												-0.076362835						
51	Arsenic	U	30	2												-0.076362835						
52	Arsenic	U	31	2												-0.076362835						
53	Arsenic	U	32	11												0.007416224						
54	Arsenic	U	33	13												0.036802012						
55	Arsenic	U	34	12												0.018361986						
56	Arsenic	U	35	12												0.018361986						
57	Arsenic	U	36	13												0.036802012						
58	Arsenic	U	37	10												0.002010125						
59	Arsenic	U	38	11												0.007416224						
60	Arsenic	U	39	11												0.007416224						
61	Arsenic	U	40	11												0.007416224						
62	Arsenic	U	41	12												0.018361986						
63	Arsenic	U	42	9.9												0.001698945						
105	Arsenic	U	43	9.96												0.00188146						
107	Arsenic	U	44	2												-0.076362835						
108	Arsenic	U	45	2												-0.076362835						
109	Arsenic	U	46	2												-0.076362835						
110	Arsenic	U	47	2												-0.076362835						
119	Arsenic	U	48	8												-1.48567E-06						

Arsenic urban robust percentile calculation spreadsheet

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
48	11.9	8	-3.9	-7.96	-0.16583	7.668333333	8.331666667	8.284753	8.37858
	7		1						
	7		1						
	7		1						
	8		0						
	8		0						
	6.2		1.8						
	5.9		2.1						
	6.8		1.2						
	6.3		1.7						
	6.9		1.1						
	6		2						
	7.9		0.1						
	6.3		1.7						
	14		-6						
	23		-15						
	9		-1						
	8		0						
	8		0						
	10		-2						
	9		-1						
	7		1						
	5.2		2.8						
	6.5		1.5						
	4.5		3.5						
	6		2						
	11.8		-3.8						
	10.9		-2.9						
	2		6						
	2		6						
	2		6						
	11		-3						
	13		-5						
	12		-4						
	12		-4						
	13		-5						
	10		-2						
	11		-3						
	11		-3						
	11		-3						
	12		-4						
	9.9		-1.9						
	9.96		-1.96						
	2		6						
	2		6						
	2		6						
	2		6						
	8		0						

Arsenic urban empirical calculation spreadsheet

XLSTAT 2013.4.06 - Quantiles estimation - on 27/08/2013 at 13:09:25
Data: Workbook = As Urban Sheet.xlsx / Sheet = Empirical / Range = Empirical!$A$1:$A$49 / 48 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

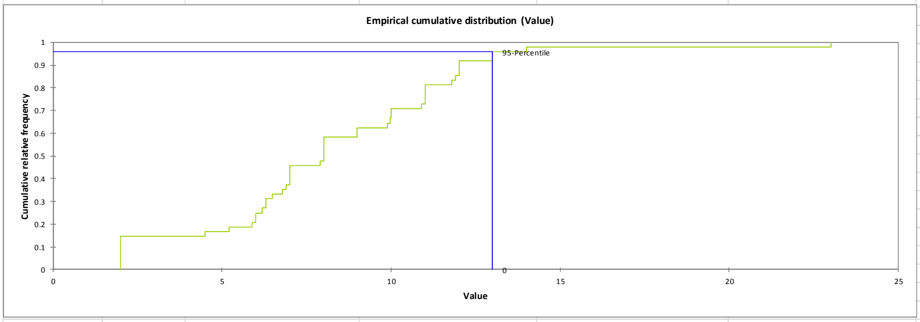
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	48	0	48	2.000	23.000	8.166	3.999

Percentile table (Weighted average at x(Np)):

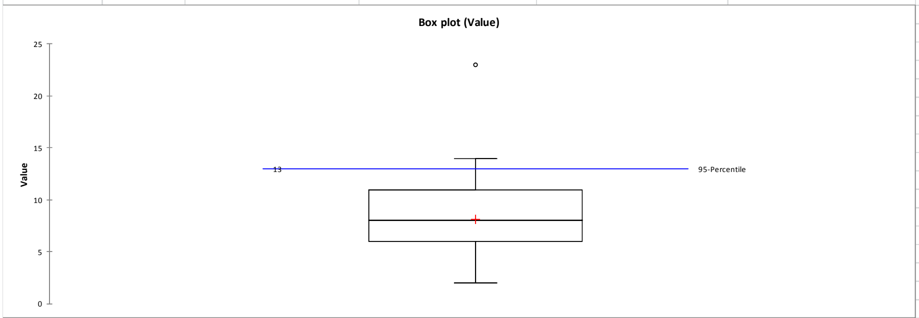
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	23.000				
99%	18.680	13.000	23.000	13.000	23.000
95%	13.000	12.000	23.000	12.000	23.000
90%	12.000	11.000	14.000	11.000	23.000
3rd Quartile 75%	11.000	9.000	12.000	9.000	12.000
Median 50%	8.000	6.800	9.900	6.800	9.900
1st Quartile 25%	6.000	2.000	6.900	2.000	6.900
10%	2.000	2.000	5.200	2.000	4.500
5%	2.000	2.000	2.000	2.000	4.500
1%	2.000	2.000	2.000	2.000	4.500
Minimum 0%	2.000				

Value of the 95-percentile: 13

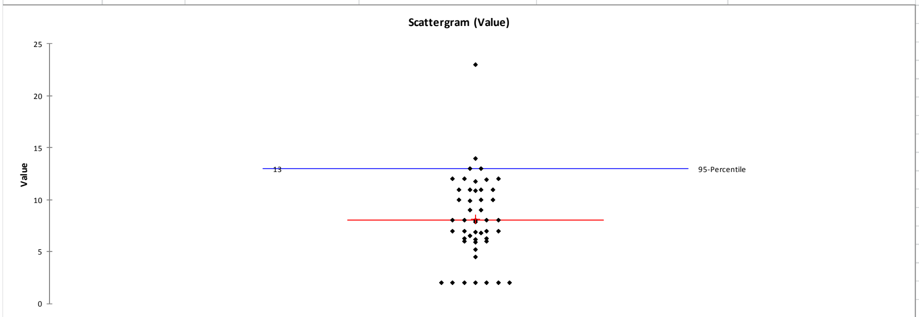
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 4 – Benzo[a]pyrene Rural Data

Full spreadsheet for benzo[a]pyrene rural, including parametric percentiles calculation on original data

ID	Contaminant	Area	Data Point No.	Value	(xi-μ)^3/Nσ^3
18	BaP	R	1	1.3	3.568119595
19	BaP	R	2	0.1	-0.000670181
20	BaP	R	3	0.1	-0.000670181
21	BaP	R	4	0.1	-0.000670181
22	BaP	R	5	0.1	-0.000670181
23	BaP	R	6	0.1	-0.000670181
44	BaP	R	7	0.71	0.395040229
45	BaP	R	8	0.43	0.045413967
46	BaP	R	9	0.36	0.018094878
64	BaP	R	10	0.1	-0.000670181
65	BaP	R	11	0.1	-0.000670181
66	BaP	R	12	0.1	-0.000670181
67	BaP	R	13	0.1	-0.000670181
68	BaP	R	14	0.1	-0.000670181
69	BaP	R	15	0.1	-0.000670181
70	BaP	R	16	0.1	-0.000670181
71	BaP	R	17	0.1	-0.000670181
72	BaP	R	18	0.1	-0.000670181
73	BaP	R	19	0.1	-0.000670181
74	BaP	R	20	0.1	-0.000670181
75	BaP	R	21	0.2	0.00010463
76	BaP	R	22	0.1	-0.000670181
77	BaP	R	23	0.1	-0.000670181
78	BaP	R	24	0.1	-0.000670181
79	BaP	R	25	0.1	-0.000670181
80	BaP	R	26	0.1	-0.000670181
81	BaP	R	27	0.2	0.00010463
83	BaP	R	28	0.1	-0.000670181
84	BaP	R	29	0.1	-0.000670181
85	BaP	R	30	0.1	-0.000670181
86	BaP	R	31	0.1	-0.000670181
87	BaP	R	32	0.1	-0.000670181
88	BaP	R	33	0.1	-0.000670181
89	BaP	R	34	0.1	-0.000670181
93	BaP	R	35	0.1	-0.000670181
94	BaP	R	36	0.1	-0.000670181
97	BaP	R	37	0.1	-0.000670181
98	BaP	R	38	0.1	-0.000670181
103	BaP	R	39	0.1	-0.000670181
104	BaP	R	40	0.1	-0.000670181

ID	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
18	40	0.165	0.217185	0.1	0.1	0.1	0.1	0.1	0.2	0.444	1.3	4.004091786	4.322906112	1	non-Gaussian	0.376695	0.511305

Benzo[a]pyrene rural robust percentile calculation spreadsheet on original data

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
40	1.3	0.1	-1.2	-2.6	-0.065	-0.03	0.23	0.209857	0.250143
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.71		-0.61						
	0.43		-0.33						
	0.36		-0.26						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.2		-0.1						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.2		-0.1						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						
	0.1		0						

Benzo[a]pyrene rural empirical calculation spreadsheet for original data

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 16:06:00
Data: Workbook = BaP Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$41 / 40 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

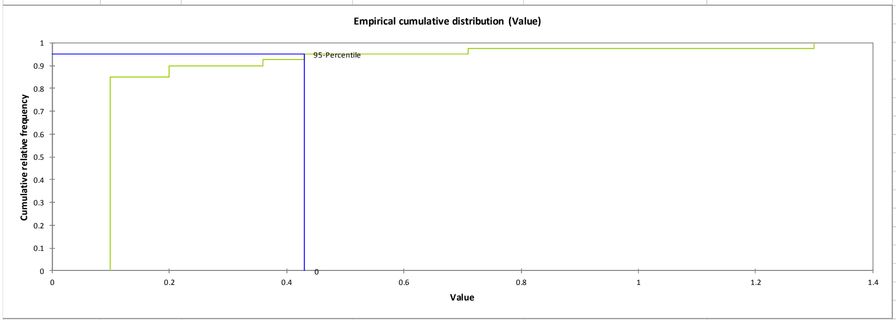
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	40	0	40	0.100	1.300	0.165	0.217

Percentile table (Weighted average at x(Np)):

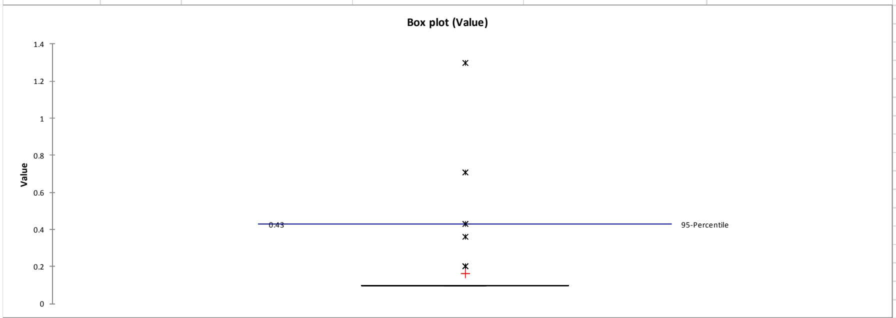
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	1.300				
99%	1.064	0.430	1.300	0.430	1.300
95%	0.430	0.200	1.300	0.200	1.300
90%	0.200	0.100	1.300	0.100	1.300
3rd Quartile 75%	0.100	0.100	0.200	0.100	0.200
Median 50%	0.100	0.100	0.100	0.100	0.100
1st Quartile 25%	0.100	0.100	0.100	0.100	0.100
10%	0.100	0.100	0.100	0.100	0.100
5%	0.100	0.100	0.100	0.100	0.200
1%	0.100	0.100	0.100	0.100	0.200
Minimum 0%	0.100				

Value of the 95-percentile: 0.43

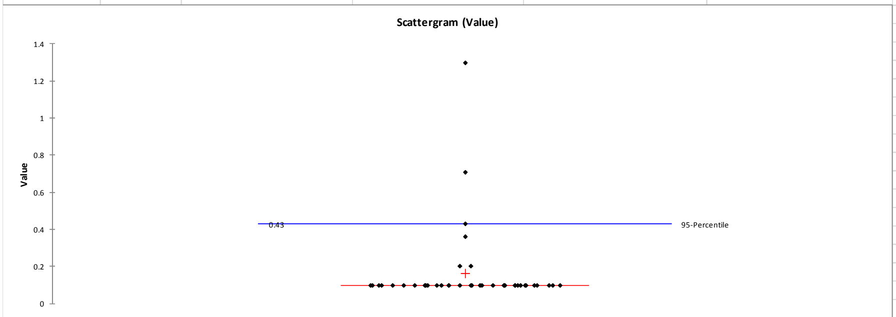
Empirical cumulative distribution:



Box plots:



Scattergrams:



Full spreadsheet for benzo[a]pyrene rural, log transform data

Log Value	No. of data points	Log mean	Log std. dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification
0.113943352	40	-0.906074184	0.253418731	-1	-1	-1	-1	-1	-0.698970004	-0.35564205	0.1139434	1.630222135	2.674596369	2.887553435	1	non-Gaussian
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-0.148741651												0.667242515				
-0.366531544												0.241268813				
-0.443697499												0.1518489				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-0.698970004												0.013645532				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-0.698970004												0.013645532				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				
-1												-0.001272855				

Full spreadsheet for benzo[a]pyrene rural, Box-Cox transform data

Box-Cox Value	No. of data points	Box-Cox Mean	Box-Cox Std. Dev.	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification
1.140175425	40	0.372117637	0.164950541	0.31622777	0.316227766	0.316227766	0.316228	0.316227766	0.447213595	0.665087409	1.1401754	2.523832131	3.268660313	3.528918017	1	non-Gaussian
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.842614977												0.580163351				
0.655743852												0.127091903				
0.6												0.065919046				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.447213595												0.002359003				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.447213595												0.002359003				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				
0.316227766												-0.000972474				

Appendix 5 – Benzo[a]pyrene Urban Data

Full spreadsheet for benzo[a]pyrene urban original data

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification
1	BaP	U	1	4.5	30	0.796666667	0.909522	0.1	0.1	0.15	0.5	1.1	1.3375	2.175	4.5	2.250182975	2.370550404	2.627457344	0.353535354	non-Gaussian
2	BaP	U	2	0.1												-0.014980093				
3	BaP	U	3	0.1												-0.014980093				
4	BaP	U	4	0.1												-0.014980093				
5	BaP	U	5	0.1												-0.014980093				
6	BaP	U	6	0.1												-0.014980093				
9	BaP	U	7	0.8												1.64088E-09				
10	BaP	U	8	0.5												-0.001156766				
11	BaP	U	9	0.5												-0.001156766				
12	BaP	U	10	0.3												-0.005427933				
13	BaP	U	11	0.5												-0.001156766				
14	BaP	U	12	0.5												-0.001156766				
24	BaP	U	13	1.4												0.009729967				
25	BaP	U	14	0.3												-0.005427933				
26	BaP	U	15	0.4												-0.002765136				
27	BaP	U	16	0.1												-0.014980093				
28	BaP	U	17	0.1												-0.014980093				
47	BaP	U	18	0.9												4.88833E-05				
48	BaP	U	19	0.5												-0.001156766				
53	BaP	U	20	0.1												-0.014980093				
54	BaP	U	21	0.5												-0.001156766				
55	BaP	U	22	0.3												-0.005427933				
56	BaP	U	23	1.2												0.002906911				
57	BaP	U	24	2.4												0.182604252				
58	BaP	U	25	1												0.000372448				
59	BaP	U	26	1.1												0.001236516				
60	BaP	U	27	1.1												0.001236516				
61	BaP	U	28	1.9												0.059505847				
62	BaP	U	29	1.2												0.002906911				
63	BaP	U	30	1.3												0.005649454				

Full spreadsheet for benzo[a]pyrene urban log transform data, including parametric percentiles calculation

Log Value	No. of data points	Log mean	Log std. dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
0.653212514	30	-0.336734237	0.485347985	-1	-1	-0.880719686	-0.30103	0.041392685	0.126012609	0.334555303	0.6532125	0.28284894	-0.072598819	-0.080466672	-0.241495875	Symmetrical	0.160879	0.508232
-1												-0.08507105						
-1												-0.08507105						
-1												-0.08507105						
-1												-0.08507105						
-1												-0.08507105						
-0.096910013												0.004021594						
-0.301029996												1.32702E-05						
-0.301029996												1.32702E-05						
-0.522878745												-0.001880484						
-0.301029996												1.32702E-05						
-0.301029996												1.32702E-05						
0.146128036												0.032823801						
-0.522878745												-0.001880484						
-0.397940009												-6.68492E-05						
-1												-0.08507105						
-1												-0.08507105						
-0.045757491												0.007182802						
-0.301029996												1.32702E-05						
-1												-0.08507105						
-0.301029996												1.32702E-05						
-0.522878745												-0.001880484						
0.079181246												0.020976562						
0.380211242												0.107442831						
0												0.011132197						
0.041392685												0.015762745						
0.041392685												0.015762745						
0.278753601												0.067979424						
0.079181246												0.020976562						
0.113943352												0.026688056						

Benzo[a]pyrene urban robust percentile calculation on log transform data

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
30	0.653213	-0.30103	-0.954242509	1.071127245	0.035704	-0.229621513	-0.372438479	-0.2424	-0.21685
	-1		0.698970004						
	-1		0.698970004						
	-1		0.698970004						
	-1		0.698970004						
	-1		0.698970004						
	-0.09691		-0.204119983						
	-0.30103		0						
	-0.30103		0						
	-0.52288		0.22184875						
	-0.30103		0						
	-0.30103		0						
	0.146128		-0.447158031						
	-0.52288		0.22184875						
	-0.39794		0.096910013						
	-1		0.698970004						
	-1		0.698970004						
	-0.04576		-0.255272505						
	-0.30103		0						
	-1		0.698970004						
	-0.30103		0						
	-0.52288		0.22184875						
	0.079181		-0.380211242						
	0.380211		-0.681241237						
	0		-0.301029996						
	0.041393		-0.342422681						
	0.041393		-0.342422681						
	0.278754		-0.579783597						
	0.079181		-0.380211242						
	0.113943		-0.414973348						

Benzo[a]pyrene urban empirical calculation on log transform data

XLSTAT 2013.4.06 - Quantiles estimation - on 27/08/2013 at 12:20:36
Data: Workbook = BaP Urban Sheet.xlsx / Sheet = Empirical / Range = Empirical!$A$1:$A$31 / 30 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

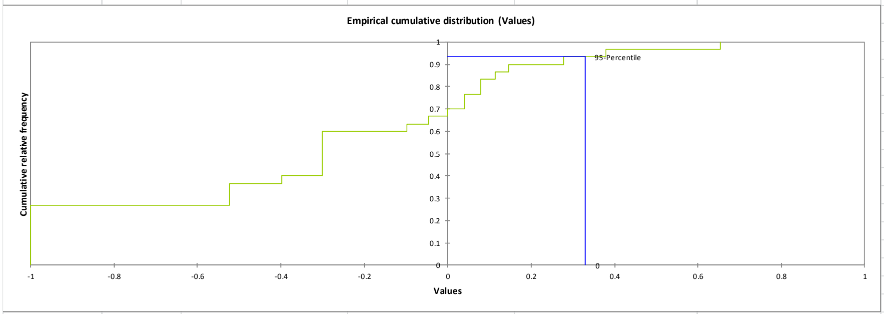
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Values	30	0	30	-1.000	0.653	-0.337	0.485

Percentile table (Weighted average at x(Np)):

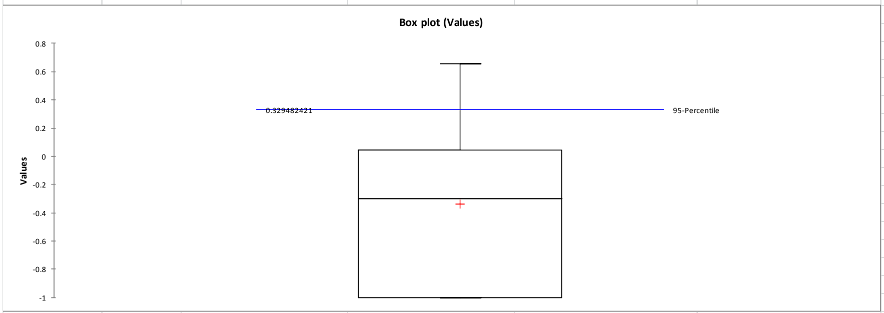
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	0.653				
99%	0.571	0.380	0.653	0.279	0.653
95%	0.329	0.114	0.653	0.114	0.653
90%	0.146	0.079	0.653	0.079	0.653
3rd Quartile 75%	0.041	-0.301	0.146	-0.301	0.146
Median 50%	-0.301	-0.523	-0.046	-0.523	0.000
1st Quartile 25%	-1.000	-1.000	-0.398	-1.000	-0.301
10%	-1.000	-1.000	-1.000	-1.000	-0.523
5%	-1.000	-1.000	-1.000	-1.000	-0.523
1%	-1.000	-1.000	-1.000	-1.000	-0.523
Minimum 0%	-1.000				

Value of the 95-percentile: 0.329482421332218

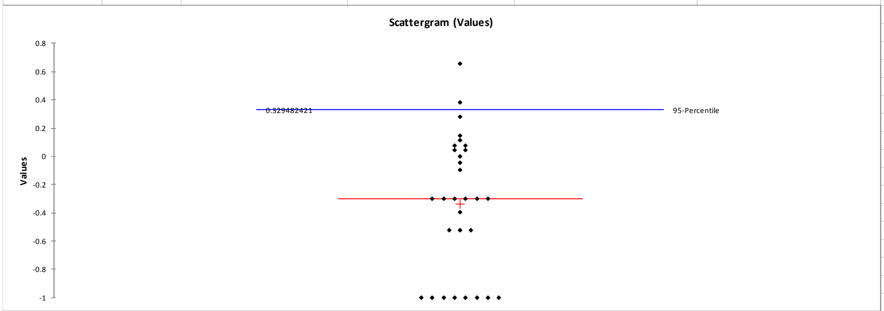
Empirical cumulative distribution:



Box plots:



Scattergrams:



Benzo[a]pyrene urban back transformation of data

Percentile	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	R H
95	0.329482	0.113943	0.653213	0.3345553	0.160544	0.508566	-0.22962	-0.2424	-0.21685
Back transformation of data									
Percentile	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	R H
95	2.135416	1.3	4.5	2.16050514	1.447252	3.225272	0.589357	0.572272	0.606953

Appendix 6 – Cadmium Rural Data

Full spreadsheet for cadmium rural, including parametric percentiles calculation

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
18	Cadmium	R	1	0.46	70	0.516571429	0.24199	0.02	0.2	0.5	0.5	0.5	0.72125	0.964	1.5	-0.000182516	1.18677331	1.239383892	-0.151079137	Gaussian	0.907311	1.020689
19	Cadmium	R	2	0.5												-4.58765E-06						
20	Cadmium	R	3	0.4												-0.001596934						
21	Cadmium	R	4	0.4												-0.001596934						
22	Cadmium	R	5	0.5												-4.58765E-06						
23	Cadmium	R	6	0.5												-4.58765E-06						
32	Cadmium	R	7	0.19												-0.03511115						
33	Cadmium	R	8	0.19												-0.03511115						
35	Cadmium	R	9	0.2												-0.031983474						
36	Cadmium	R	10	0.2												-0.031983474						
37	Cadmium	R	11	0.3												-0.010240331						
38	Cadmium	R	12	0.5												-4.58765E-06						
39	Cadmium	R	13	0.4												-0.001596934						
40	Cadmium	R	14	1.1												0.200204379						
41	Cadmium	R	15	0.2												-0.031983474						
42	Cadmium	R	16	0.2												-0.031983474						
43	Cadmium	R	17	0.2												-0.031983474						
44	Cadmium	R	18	0.16												-0.045703569						
45	Cadmium	R	19	0.14												-0.053833497						
46	Cadmium	R	20	0.15												-0.049657669						
64	Cadmium	R	21	0.5												-4.58765E-06						
65	Cadmium	R	22	0.5												-4.58765E-06						
66	Cadmium	R	23	0.5												-4.58765E-06						
67	Cadmium	R	24	0.5												-4.58765E-06						
68	Cadmium	R	25	0.5												-4.58765E-06						
69	Cadmium	R	26	0.5												-4.58765E-06						
70	Cadmium	R	27	0.5												-4.58765E-06						
71	Cadmium	R	28	0.5												-4.58765E-06						
72	Cadmium	R	29	0.5												-4.58765E-06						
73	Cadmium	R	30	0.5												-4.58765E-06						
74	Cadmium	R	31	0.5												-4.58765E-06						
75	Cadmium	R	32	0.5												-4.58765E-06						
76	Cadmium	R	33	0.5												-4.58765E-06						
77	Cadmium	R	34	0.5												-4.58765E-06						
78	Cadmium	R	35	0.5												-4.58765E-06						
79	Cadmium	R	36	0.5												-4.58765E-06						
80	Cadmium	R	37	0.5												-4.58765E-06						
81	Cadmium	R	38	0.5												-4.58765E-06						
82	Cadmium	R	39	0.5												-4.58765E-06						
83	Cadmium	R	40	0.5												-4.58765E-06						
84	Cadmium	R	41	0.6												0.000585403						
85	Cadmium	R	42	0.74												0.01124415						
86	Cadmium	R	43	0.92												0.066192754						
87	Cadmium	R	44	0.68												0.004400423						
88	Cadmium	R	45	0.5												-4.58765E-06						
89	Cadmium	R	46	0.5												-4.58765E-06						
90	Cadmium	R	47	0.71												0.007295791						
91	Cadmium	R	48	0.81												0.02546934						
92	Cadmium	R	49	0.5												-4.58765E-06						
93	Cadmium	R	50	0.59												0.000399122						
94	Cadmium	R	51	0.5												-4.58765E-06						
95	Cadmium	R	52	0.5												-4.58765E-06						
96	Cadmium	R	53	0.5												-4.58765E-06						
97	Cadmium	R	54	0.56												8.25728E-05						
98	Cadmium	R	55	0.5												-4.58765E-06						
99	Cadmium	R	56	0.5												-4.58765E-06						
100	Cadmium	R	57	0.5												-4.58765E-06						
101	Cadmium	R	58	0.5												-4.58765E-06						
102	Cadmium	R	59	0.5												-4.58765E-06						
103	Cadmium	R	60	0.5												-4.58765E-06						
104	Cadmium	R	61	0.64												0.001895648						
106	Cadmium	R	62	0.02												-0.123439932						
111	Cadmium	R	63	0.5												-4.58765E-06						
112	Cadmium	R	64	1												0.113895728						
113	Cadmium	R	65	0.5												-4.58765E-06						
114	Cadmium	R	66	0.6												0.000585403						
115	Cadmium	R	67	0.9												0.056828201						
116	Cadmium	R	68	1.5												0.958824134						
117	Cadmium	R	69	0.9												0.056828201						
118	Cadmium	R	70	1.1												0.200204379						

Cadmium rural robust percentile calculation

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
70	0.46	0.5	0.04	-1.16	-0.01657	0.466857143	0.533142857	0.529261	0.537025
	0.5		0						
	0.4		0.1						
	0.4		0.1						
	0.5		0						
	0.5		0						
	0.19		0.31						
	0.19		0.31						
	0.2		0.3						
	0.2		0.3						
	0.3		0.2						
	0.5		0						
	0.4		0.1						
	1.1		-0.6						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.16		0.34						
	0.14		0.36						
	0.15		0.35						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.6		-0.1						
	0.74		-0.24						
	0.92		-0.42						
	0.68		-0.18						
	0.5		0						
	0.5		0						
	0.71		-0.21						
	0.81		-0.31						
	0.5		0						
	0.59		-0.09						
	0.5		0						
	0.5		0						
	0.5		0						
	0.56		-0.06						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.64		-0.14						
	0.02		0.48						
	0.5		0						
	1		-0.5						
	0.5		0						
	0.6		-0.1						
	0.9		-0.4						
	1.5		-1						
	0.9		-0.4						
	1.1		-0.6						

Cadmium rural empirical percentile calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 16:16:01
Data: Workbook = Cd Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$71 / 70 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

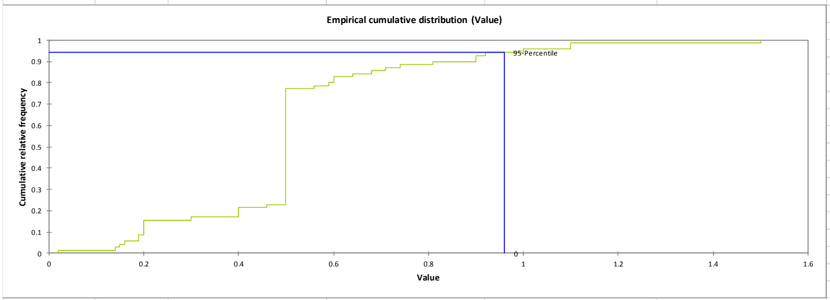
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	70	0	70	0.020	1.500	0.517	0.242

Percentile table (Weighted average at x(Np)):

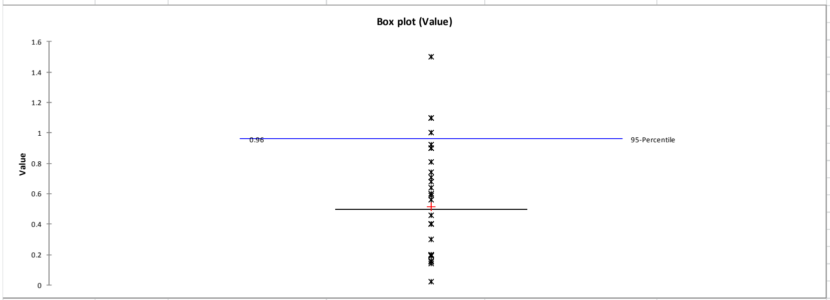
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	1.500				
99%	1.220	1.100	1.500	1.100	1.500
95%	0.960	0.810	1.500	0.740	1.500
90%	0.810	0.600	1.100	0.600	1.100
3rd Quartile 75%	0.500	0.500	0.680	0.500	0.680
Median 50%	0.500	0.500	0.500	0.500	0.500
1st Quartile 25%	0.500	0.200	0.500	0.200	0.500
10%	0.200	0.140	0.300	0.140	0.300
5%	0.155	0.020	0.200	0.020	0.190
1%	0.104	0.020	0.140	0.020	0.140
Minimum 0%	0.020				

Value of the 95-percentile: 0.96

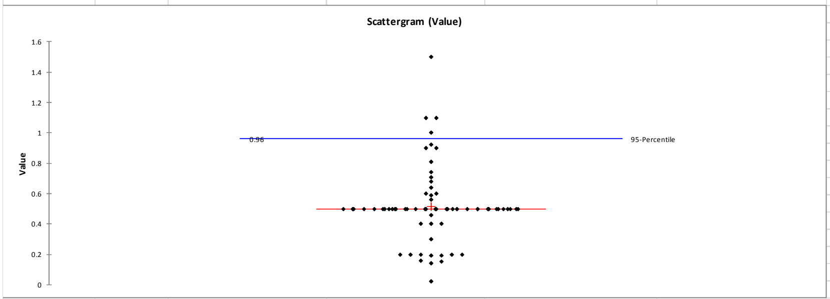
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 7 – Cadmium Urban Data

Full spreadsheet for cadmium urban, including parametric percentile calculation

ID	Contaminant	Area	Data Point No.	Value	(xi-μ)^3/Nσ^3
1	Cadmium	U	1	0.4	-0.008243408
2	Cadmium	U	2	1.2	0.085387266
3	Cadmium	U	3	1.1	0.046677374
4	Cadmium	U	4	1	0.021895655
5	Cadmium	U	5	1.1	0.046677374
6	Cadmium	U	6	1.1	0.046677374
7	Cadmium	U	7	0.5	-0.001803082
8	Cadmium	U	8	0.5	-0.001803082
9	Cadmium	U	9	0.6	-7.10541E-05
10	Cadmium	U	10	0.6	-7.10541E-05
11	Cadmium	U	11	0.7	5.87551E-05
12	Cadmium	U	12	0.5	-0.001803082
13	Cadmium	U	13	0.6	-7.10541E-05
14	Cadmium	U	14	0.6	-7.10541E-05
15	Cadmium	U	15	0.35	-0.014199831
16	Cadmium	U	16	0.4	-0.008243408
17	Cadmium	U	17	1.1	0.046677374
24	Cadmium	U	18	0.9	0.007936031
25	Cadmium	U	19	0.8	0.001692424
26	Cadmium	U	20	0.9	0.007936031
27	Cadmium	U	21	0.8	0.001692424
28	Cadmium	U	22	1	0.021895655
29	Cadmium	U	23	0.2	-0.047673264
30	Cadmium	U	24	0.2	-0.047673264
31	Cadmium	U	25	0.2	-0.047673264
34	Cadmium	U	26	0.6	-7.10541E-05
47	Cadmium	U	27	0.5	-0.001803082
48	Cadmium	U	28	0.5	-0.001803082
50	Cadmium	U	29	0.2	-0.047673264
51	Cadmium	U	30	0.2	-0.047673264
52	Cadmium	U	31	0.2	-0.047673264
53	Cadmium	U	32	0.9	0.007936031
54	Cadmium	U	33	1	0.021895655
55	Cadmium	U	34	1.1	0.046677374
56	Cadmium	U	35	1	0.021895655
57	Cadmium	U	36	1	0.021895655
58	Cadmium	U	37	0.8	0.001692424
59	Cadmium	U	38	1	0.021895655
60	Cadmium	U	39	1	0.021895655
61	Cadmium	U	40	0.9	0.007936031
62	Cadmium	U	41	1	0.021895655
63	Cadmium	U	42	0.9	0.007936031
105	Cadmium	U	43	0.126	-0.075159854
107	Cadmium	U	44	0.2	-0.047673264
108	Cadmium	U	45	0.2	-0.047673264
109	Cadmium	U	46	0.2	-0.047673264
110	Cadmium	U	47	0.2	-0.047673264
119	Cadmium	U	48	0.2	-0.047673264

ID	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
1	48	0.6515833	0.342688	0.126	0.2	0.3125	0.6	1	1.0125	1.1	1.2	-0.100867533	-0.107492505	0.015384615	Symmetrical	1.003055	1.196945

Cadmium urban robust percentile calculation

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
48	0.4	0.6	0.2	-2.476	-0.05158	0.496833333	0.703166667	0.688574	0.717759
	1.2		-0.6						
	1.1		-0.5						
	1		-0.4						
	1.1		-0.5						
	1.1		-0.5						
	0.5		0.1						
	0.5		0.1						
	0.6		0						
	0.6		0						
	0.7		-0.1						
	0.5		0.1						
	0.6		0						
	0.6		0						
	0.35		0.25						
	0.4		0.2						
	1.1		-0.5						
	0.9		-0.3						
	0.8		-0.2						
	0.9		-0.3						
	0.8		-0.2						
	1		-0.4						
	0.2		0.4						
	0.2		0.4						
	0.2		0.4						
	0.6		0						
	0.5		0.1						
	0.5		0.1						
	0.2		0.4						
	0.2		0.4						
	0.2		0.4						
	0.9		-0.3						
	1		-0.4						
	1.1		-0.5						
	1		-0.4						
	1		-0.4						
	0.8		-0.2						
	1		-0.4						
	1		-0.4						
	0.9		-0.3						
	1		-0.4						
	0.9		-0.3						
	0.126		0.474						
	0.2		0.4						
	0.2		0.4						
	0.2		0.4						
	0.2		0.4						
	0.2		0.4						

Cadmium urban empirical percentiles calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 27/08/2013 at 13:18:01
Data: Workbook = Cd Urban Sheet.xlsx / Sheet = Empirical / Range = Empirical!$A$1:$A$49 / 48 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

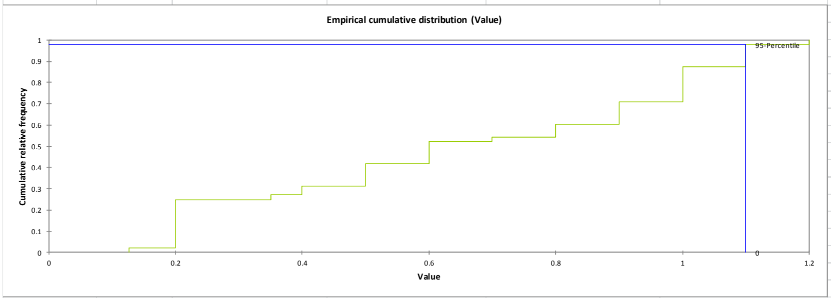
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	48	0	48	0.126	1.200	0.652	0.343

Percentile table (Weighted average at x(Np)):

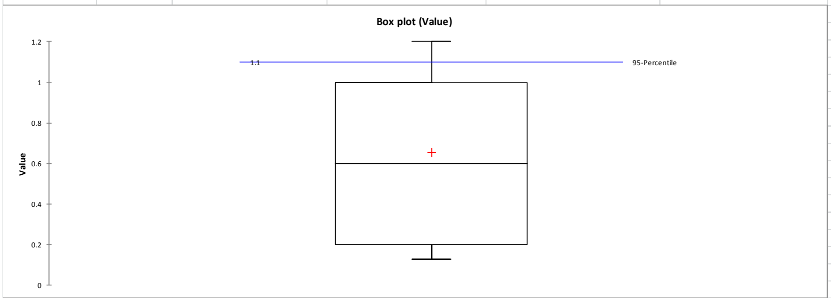
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	1.200				
99%	1.152	1.100	1.200	1.100	1.200
95%	1.100	1.100	1.200	1.000	1.200
90%	1.100	1.000	1.100	1.000	1.200
3rd Quartile 75%	1.000	0.900	1.000	0.900	1.000
Median 50%	0.600	0.500	0.900	0.500	0.900
1st Quartile 25%	0.200	0.200	0.500	0.200	0.500
10%	0.200	0.126	0.200	0.126	0.200
5%	0.200	0.126	0.200	0.126	0.350
1%	0.162	0.126	0.200	0.126	0.200
Minimum 0%	0.126				

Value of the 95-percentile: 1.1

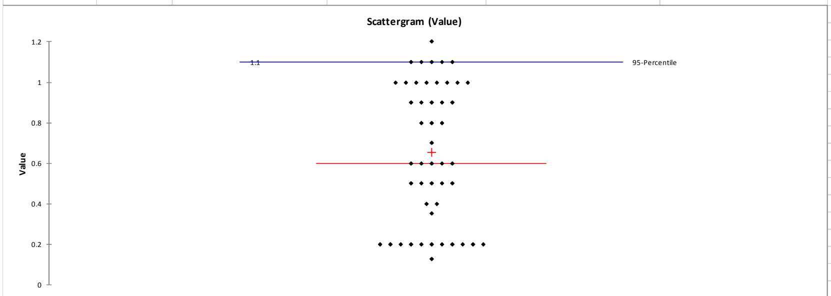
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 8 – Copper Rural Data

Full spreadsheet for copper rural original data

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification
18	Copper	R	1	53	70	25.52714286	38.88177	8.7	13.625	15	18.5	24	27.375	45.8	330	0.005039359	6.925839501	7.232867339	0.290909091	non-Gaussian
19	Copper	R	2	19												-6.75823E-05				
20	Copper	R	3	14												-0.000372245				
21	Copper	R	4	33												0.00010142				
22	Copper	R	5	21												-2.25495E-05				
23	Copper	R	6	21												-2.25495E-05				
32	Copper	R	7	18												-0.000103646				
33	Copper	R	8	17												-0.000150687				
35	Copper	R	9	27												7.76508E-07				
36	Copper	R	10	25												-3.55999E-08				
37	Copper	R	11	28												3.67503E-06				
38	Copper	R	12	28												3.67503E-06				
39	Copper	R	13	26												2.56953E-08				
40	Copper	R	14	37												0.000367011				
41	Copper	R	15	24												-8.65571E-07				
42	Copper	R	16	11.4												-0.000685215				
43	Copper	R	17	11.9												-0.000615004				
44	Copper	R	18	22												-1.06643E-05				
45	Copper	R	19	22												-1.06643E-05				
46	Copper	R	20	26												2.56953E-08				
64	Copper	R	21	14												-0.000372245				
65	Copper	R	22	11												-0.000745083				
66	Copper	R	23	14												-0.000372245				
67	Copper	R	24	24												-8.65571E-07				
68	Copper	R	25	19												-6.75823E-05				
69	Copper	R	26	21												-2.25495E-05				
70	Copper	R	27	21												-2.25495E-05				
71	Copper	R	28	20												-4.10361E-05				
72	Copper	R	29	20												-4.10361E-05				
73	Copper	R	30	19												-6.75823E-05				
74	Copper	R	31	23												-3.92241E-06				
75	Copper	R	32	18												-0.000103646				
76	Copper	R	33	30												2.1748E-05				
77	Copper	R	34	19												-6.75823E-05				
78	Copper	R	35	27												7.76508E-07				
79	Copper	R	36	20												-4.10361E-05				
80	Copper	R	37	18												-0.000103646				
81	Copper	R	38	15												-0.000283528				
82	Copper	R	39	12												-0.000601564				
83	Copper	R	40	16												-0.000210161				
84	Copper	R	41	15												-0.000283528				
85	Copper	R	42	16												-0.000210161				
86	Copper	R	43	18												-0.000103646				
87	Copper	R	44	15												-0.000283528				
88	Copper	R	45	15												-0.000283528				
89	Copper	R	46	80												0.039282968				
90	Copper	R	47	14												-0.000372245				
91	Copper	R	48	15												-0.000283528				
92	Copper	R	49	13												-0.000477771				
93	Copper	R	50	14												-0.000372245				
94	Copper	R	51	9.6												-0.000981923				
95	Copper	R	52	16												-0.000210161				
96	Copper	R	53	20												-4.10361E-05				
97	Copper	R	54	10												-0.000909784				
98	Copper	R	55	14												-0.000372245				
99	Copper	R	56	8.7												-0.001157964				
100	Copper	R	57	16												-0.000210161				
101	Copper	R	58	16												-0.000210161				
102	Copper	R	59	16												-0.000210161				
103	Copper	R	60	17												-0.000150687				
104	Copper	R	61	26												2.56953E-08				
106	Copper	R	62	16.3												-0.000190926				
111	Copper	R	63	21												-2.25495E-05				
112	Copper	R	64	78												0.035113002				
113	Copper	R	65	15												-0.000283528				
114	Copper	R	66	10												-0.000909784				
115	Copper	R	67	25												-3.55999E-08				
116	Copper	R	68	330												6.859770377				
117	Copper	R	69	18												-0.000103646				
118	Copper	R	70	24												-8.65571E-07				

Full spreadsheet for copper rural log transform data including parametric percentiles calculation

Log Value	No. of data points	Log mean	Log Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
1.72427587	70	1.298761942	0.231097998	0.93951925	1.134058779	1.176091259	1.267013	1.380211242	1.437286614	1.654042504	2.5185139	0.089177388	2.485498473	2.595682549	0.123073423	Gaussian	1.599905	1.70818
1.278753601												-9.27143E-06						
1.146128036												-0.004115916						
1.51851394												0.012283211						
1.322219295												1.494E-05						
1.322219295												1.494E-05						
1.255272505												-9.52061E-05						
1.230448921												-0.000368998						
1.431363764												0.002698745						
1.397940009												0.001129173						
1.447158031												0.003782516						
1.447158031												0.003782516						
1.414973348												0.001816602						
1.568201724												0.022641154						
1.380211242												0.000625425						
1.056904851												-0.016375323						
1.075546961												-0.012873108						
1.342422681												9.63356E-05						
1.342422681												9.63356E-05						
1.414973348												0.001816602						
1.146128036												-0.004115916						
1.041392685												-0.019732559						
1.146128036												-0.004115916						
1.380211242												0.000625425						
1.278753601												-9.27143E-06						
1.322219295												1.494E-05						
1.322219295												1.494E-05						
1.301029996												1.35043E-08						
1.301029996												1.35043E-08						
1.278753601												-9.27143E-06						
1.361727836												0.000288955						
1.255272505												-9.52061E-05						
1.477121255												0.00656751						
1.278753601												-9.27143E-06						
1.431363764												0.002698745						
1.301029996												1.35043E-08						
1.255272505												-9.52061E-05						
1.176091259												-0.002136662						
1.079181246												-0.012254508						
1.204119983												-0.000981216						
1.176091259												-0.002136662						
1.204119983												-0.000981216						
1.255272505												-9.52061E-05						
1.176091259												-0.002136662						
1.176091259												-0.002136662						
1.903089987												0.255465145						
1.146128036												-0.004115916						
1.176091259												-0.002136662						
1.113943352												-0.007307189						
1.146128036												-0.004115916						
0.982271233												-0.036694106						
1.204119983												-0.000981216						
1.301029996												1.35043E-08						
1												-0.030866631						
1.146128036												-0.004115916						
0.939519253												-0.053663282						
1.204119983												-0.000981216						
1.204119983												-0.000981216						
1.204119983												-0.000981216						
1.230448921												-0.000368998						
1.414973348												0.001816602						
1.212187604												-0.000751071						
1.322219295												1.494E-05						
1.892094603												0.241773208						
1.176091259												-0.002136662						
1												-0.030866631						
1.397940009												0.001129173						
2.51851394												2.10052581						
1.255272505												-9.52061E-05						
1.380211242												0.000625425						

Copper rural robust percentiles calculation

No. of values	Log Value	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
70	1.72427587	1.267013	-0.457262817	-2.22242226	-0.03175	1.203515274	1.330510832	1.323073	1.337948
	1.278753601		-0.011740548						
	1.146128036		0.120885017						
	1.51851394		-0.251500887						
	1.322219295		-0.055206242						
	1.322219295		-0.055206242						
	1.255272505		0.011740548						
	1.230448921		0.036564132						
	1.431363764		-0.164350711						
	1.397940009		-0.130926956						
	1.447158031		-0.180144978						
	1.447158031		-0.180144978						
	1.414973348		-0.147960295						
	1.568201724		-0.301188671						
	1.380211242		-0.113198189						
	1.056904851		0.210108202						
	1.075546961		0.191466092						
	1.342422681		-0.075409628						
	1.342422681		-0.075409628						
	1.414973348		-0.147960295						
	1.146128036		0.120885017						
	1.041392685		0.225620368						
	1.146128036		0.120885017						
	1.380211242		-0.113198189						
	1.278753601		-0.011740548						
	1.322219295		-0.055206242						
	1.322219295		-0.055206242						
	1.301029996		-0.034016943						
	1.301029996		-0.034016943						
	1.278753601		-0.011740548						
	1.361727836		-0.094714783						
	1.255272505		0.011740548						
	1.477121255		-0.210108202						
	1.278753601		-0.011740548						
	1.431363764		-0.164350711						
	1.301029996		-0.034016943						
	1.255272505		0.011740548						
	1.176091259		0.090921794						
	1.079181246		0.187831807						
	1.204119983		0.06289307						
	1.176091259		0.090921794						
	1.204119983		0.06289307						
	1.255272505		0.011740548						
	1.176091259		0.090921794						
	1.176091259		0.090921794						
	1.903089987		-0.636076934						
	1.146128036		0.120885017						
	1.176091259		0.090921794						
	1.113943352		0.153069701						
	1.146128036		0.120885017						
	0.982271233		0.28474182						
	1.204119983		0.06289307						
	1.301029996		-0.034016943						
	1		0.267013053						
	1.146128036		0.120885017						
	0.939519253		0.3274938						
	1.204119983		0.06289307						
	1.204119983		0.06289307						
	1.204119983		0.06289307						
	1.230448921		0.036564132						
	1.414973348		-0.147960295						
	1.212187604		0.054825449						
	1.322219295		-0.055206242						
	1.892094603		-0.62508155						
	1.176091259		0.090921794						
	1		0.267013053						
	1.397940009		-0.130926956						
	2.51851394		-1.251500887						
	1.255272505		0.011740548						
	1.380211242		-0.113198189						

Copper rural empirical percentiles calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 16:24:57
Data: Workbook = Cu Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$71 / 70 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

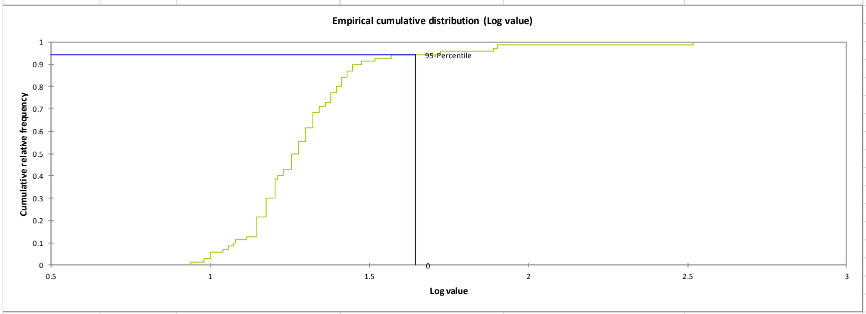
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Log value	70	0	70	0.940	2.519	1.299	0.231

Percentile table (Weighted average at x(Np)):

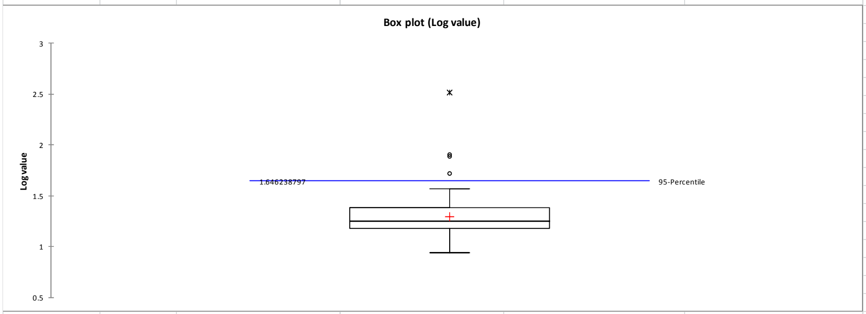
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	2.519				
99%	2.088	1.892	2.519	1.892	2.519
95%	1.646	1.447	2.519	1.447	2.519
90%	1.447	1.415	1.892	1.415	1.892
3rd Quartile 75%	1.380	1.322	1.431	1.322	1.431
Median 50%	1.255	1.204	1.301	1.204	1.322
1st Quartile 25%	1.176	1.146	1.204	1.146	1.204
10%	1.076	0.982	1.146	0.982	1.146
5%	1.000	0.940	1.076	0.940	1.041
1%	0.969	0.940	0.982	0.940	0.982
Minimum 0%	0.940				

Value of the 95-percentile: 1.64623879683389

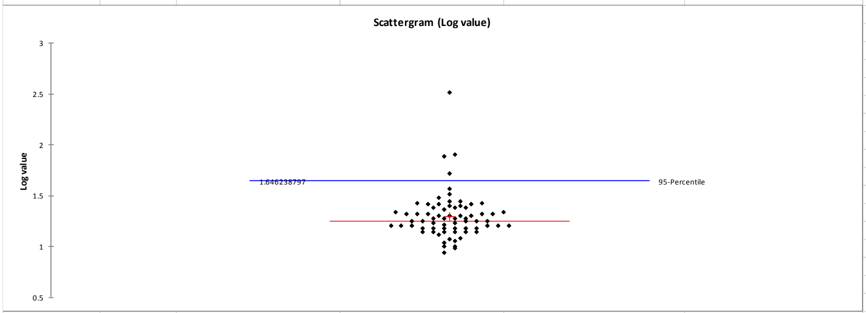
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 9 – Copper Urban Data

Full spreadsheet for copper urban, including calculation of parametric percentiles

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
1	Copper	U	1	49.5	48	23.8458333	11.30328	2	13.375	17.75	23.5	30.25	36.125	42.95	49.5	0.24356705	0.051651394	0.055043854	0.10989011	Symmetrical	39.75234	46.14766
2	Copper	U	2	18												-0.002881934						
3	Copper	U	3	17												-0.004628324						
4	Copper	U	4	16												-0.00696726						
5	Copper	U	5	22												-9.0724E-05						
6	Copper	U	6	14												-0.013768995						
7	Copper	U	7	41												0.072820503						
8	Copper	U	8	32												0.007821383						
9	Copper	U	9	18												-0.002881934						
10	Copper	U	10	17												-0.004628324						
11	Copper	U	11	20												-0.000820571						
12	Copper	U	12	17												-0.004628324						
13	Copper	U	13	20												-0.000820571						
14	Copper	U	14	18												-0.002881934						
15	Copper	U	15	29												0.001975241						
16	Copper	U	16	34												0.015103502						
17	Copper	U	17	47												0.179073889						
24	Copper	U	18	27												0.000452688						
25	Copper	U	19	28												0.001034181						
26	Copper	U	20	44												0.118097129						
27	Copper	U	21	36												0.025901229						
28	Copper	U	22	23												-8.7297E-06						
29	Copper	U	23	18												-0.002881934						
30	Copper	U	24	24												5.28587E-08						
31	Copper	U	25	9												-0.047201798						
34	Copper	U	26	17												-0.004628324						
47	Copper	U	27	29												0.001975241						
48	Copper	U	28	33												0.01106627						
50	Copper	U	29	22												-9.0724E-05						
51	Copper	U	30	9												-0.047201798						
52	Copper	U	31	20												-0.000820571						
53	Copper	U	32	25												2.21794E-05						
54	Copper	U	33	41												0.072820503						
55	Copper	U	34	31												0.005282287						
56	Copper	U	35	30												0.003362425						
57	Copper	U	36	37												0.032834834						
58	Copper	U	37	26												0.000144206						
59	Copper	U	38	25												2.21794E-05						
60	Copper	U	39	27												0.000452688						
61	Copper	U	40	27												0.000452688						
62	Copper	U	41	32												0.007821383						
63	Copper	U	42	30												0.003362425						
105	Copper	U	43	18.1												-0.002736553						
107	Copper	U	44	2												-0.150400974						
108	Copper	U	45	2												-0.150400974						
109	Copper	U	46	2												-0.150400974						
110	Copper	U	47	2												-0.150400974						
119	Copper	U	48	19												-0.001641535						

Copper urban robust percentile calculation

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
48	49.5	23.5	-26	-16.6	-0.34583	22.80833333	24.19166667	24.09383	24.2895
	18		5.5						
	17		6.5						
	16		7.5						
	22		1.5						
	14		9.5						
	41		-17.5						
	32		-8.5						
	18		5.5						
	17		6.5						
	20		3.5						
	17		6.5						
	20		3.5						
	18		5.5						
	29		-5.5						
	34		-10.5						
	47		-23.5						
	27		-3.5						
	28		-4.5						
	44		-20.5						
	36		-12.5						
	23		0.5						
	18		5.5						
	24		-0.5						
	9		14.5						
	17		6.5						
	29		-5.5						
	33		-9.5						
	22		1.5						
	9		14.5						
	20		3.5						
	25		-1.5						
	41		-17.5						
	31		-7.5						
	30		-6.5						
	37		-13.5						
	26		-2.5						
	25		-1.5						
	27		-3.5						
	27		-3.5						
	32		-8.5						
	30		-6.5						
	18.1		5.4						
	2		21.5						
	2		21.5						
	2		21.5						
	2		21.5						
	19		4.5						

Copper urban empirical percentile calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 27/08/2013 at 13:26:32
Data: Workbook = Cu Urban Sheet.xlsx / Sheet = Empirical / Range = Empirical!$A$1:$A$49 / 48 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

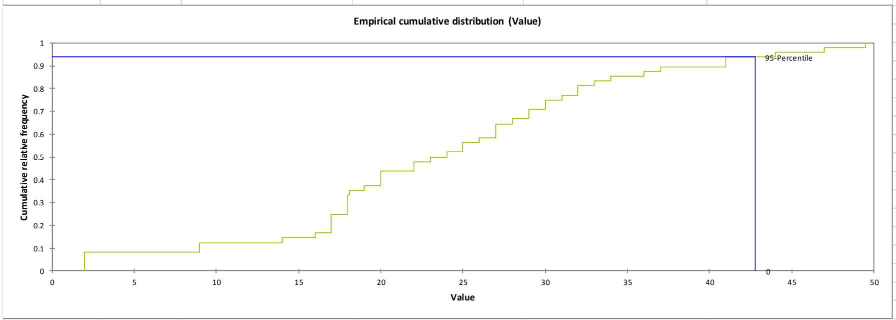
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	48	0	48	2.000	49.500	23.846	11.303

Percentile table (Weighted average at x(Np)):

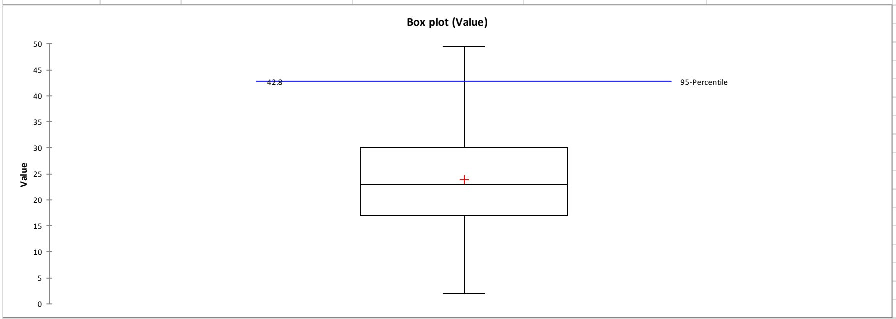
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	49.500				
99%	48.300	44.000	49.500	44.000	49.500
95%	42.800	37.000	49.500	36.000	49.500
90%	37.800	32.000	47.000	32.000	49.500
3rd Quartile 75%	30.000	27.000	36.000	27.000	36.000
Median 50%	23.000	18.100	27.000	18.100	27.000
1st Quartile 25%	17.000	9.000	19.000	9.000	19.000
10%	7.600	2.000	17.000	2.000	14.000
5%	2.000	2.000	9.000	2.000	9.000
1%	2.000	2.000	2.000	2.000	9.000
Minimum 0%	2.000				

Value of the 95-percentile: 42.8

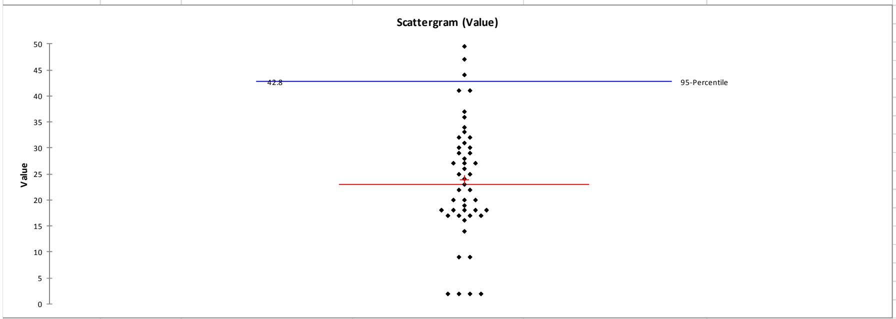
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 10 – Mercury Rural Data

Full spreadsheet for mercury rural, including calculation of parametric percentiles

ID	Contaminant	Area	Data Point No.	Value	(xi-μ)^3/Nσ^3
18	Mercury	R	1	0.54	1.44293E-06
19	Mercury	R	2	0.05	-0.005387663
20	Mercury	R	3	0.05	-0.005387663
21	Mercury	R	4	0.05	-0.005387663
22	Mercury	R	5	0.05	-0.005387663
23	Mercury	R	6	0.05	-0.005387663
32	Mercury	R	7	0.1	-0.003815853
33	Mercury	R	8	0.1	-0.003815853
35	Mercury	R	9	0.5	-6.08458E-08
36	Mercury	R	10	0.5	-6.08458E-08
37	Mercury	R	11	0.5	-6.08458E-08
38	Mercury	R	12	0.5	-6.08458E-08
39	Mercury	R	13	0.5	-6.08458E-08
40	Mercury	R	14	0.5	-6.08458E-08
41	Mercury	R	15	0.5	-6.08458E-08
42	Mercury	R	16	1	0.006485072
43	Mercury	R	17	1	0.006485072
44	Mercury	R	18	0.1	-0.003815853
45	Mercury	R	19	0.1	-0.003815853
46	Mercury	R	20	0.1	-0.003815853
64	Mercury	R	21	0.5	-6.08458E-08
65	Mercury	R	22	0.5	-6.08458E-08
66	Mercury	R	23	0.5	-6.08458E-08
67	Mercury	R	24	0.5	-6.08458E-08
68	Mercury	R	25	0.5	-6.08458E-08
69	Mercury	R	26	0.5	-6.08458E-08
70	Mercury	R	27	2.8	0.663007624
71	Mercury	R	28	0.7	0.000376885
72	Mercury	R	29	0.5	-6.08458E-08
73	Mercury	R	30	0.5	-6.08458E-08
74	Mercury	R	31	0.5	-6.08458E-08
75	Mercury	R	32	0.5	-6.08458E-08
76	Mercury	R	33	0.5	-6.08458E-08
77	Mercury	R	34	0.5	-6.08458E-08
78	Mercury	R	35	0.5	-6.08458E-08
79	Mercury	R	36	2.6	0.504002952
80	Mercury	R	37	0.9	0.00326811
81	Mercury	R	38	0.5	-6.08458E-08
82	Mercury	R	39	0.5	-6.08458E-08
83	Mercury	R	40	0.7	0.000376885
84	Mercury	R	41	0.2	-0.001650673
85	Mercury	R	42	0.2	-0.001650673
86	Mercury	R	43	0.2	-0.001650673
87	Mercury	R	44	0.2	-0.001650673
88	Mercury	R	45	0.2	-0.001650673
89	Mercury	R	46	0.2	-0.001650673
93	Mercury	R	47	0.2	-0.001650673
94	Mercury	R	48	0.2	-0.001650673
97	Mercury	R	49	0.2	-0.001650673
98	Mercury	R	50	0.2	-0.001650673
103	Mercury	R	51	0.2	-0.001650673
104	Mercury	R	52	0.2	-0.001650673
106	Mercury	R	53	0.14	-0.002805161
111	Mercury	R	54	0.3	-0.000513912
112	Mercury	R	55	0.3	-0.000513912
113	Mercury	R	56	0.3	-0.000513912
114	Mercury	R	57	0.3	-0.000513912
115	Mercury	R	58	4.1	2.554835712
116	Mercury	R	59	0.3	-0.000513912
117	Mercury	R	60	0.3	-0.000513912
118	Mercury	R	61	0.7	0.000376885

ID	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
18	61	0.5103279	0.667048	0.05	0.1	0.2	0.5	0.5	0.7	1	4.1	3.667501012	3.855020132	-0.333333333	Gaussian	0.832606	1.167394

Mercury rural robust percentiles calculation

No. of values	Value	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
61	0.54	0.5	-0.04	-0.63	-0.01033	0.479344262	0.520655738	0.518064	0.523247
	0.05		0.45						
	0.05		0.45						
	0.05		0.45						
	0.05		0.45						
	0.05		0.45						
	0.1		0.4						
	0.1		0.4						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	1		-0.5						
	1		-0.5						
	0.1		0.4						
	0.1		0.4						
	0.1		0.4						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	2.8		-2.3						
	0.7		-0.2						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	0.5		0						
	2.6		-2.1						
	0.9		-0.4						
	0.5		0						
	0.5		0						
	0.7		-0.2						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.2		0.3						
	0.14		0.36						
	0.3		0.2						
	0.3		0.2						
	0.3		0.2						
	0.3		0.2						
	4.1		-3.6						
	0.3		0.2						
	0.3		0.2						
	0.7		-0.2						

Mercury rural empirical percentiles calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 16:32:04
Data: Workbook = Hg Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$62 / 61 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

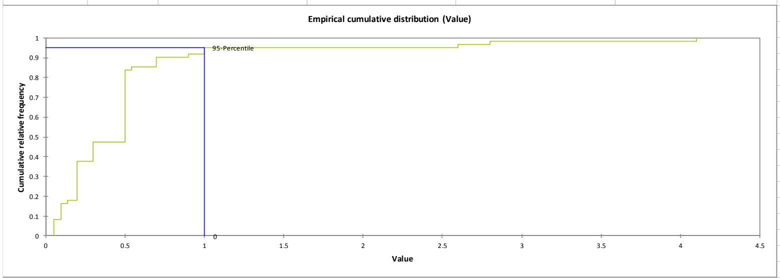
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	61	0	61	0.050	4.100	0.510	0.667

Percentile table (Weighted average at x(Np)):

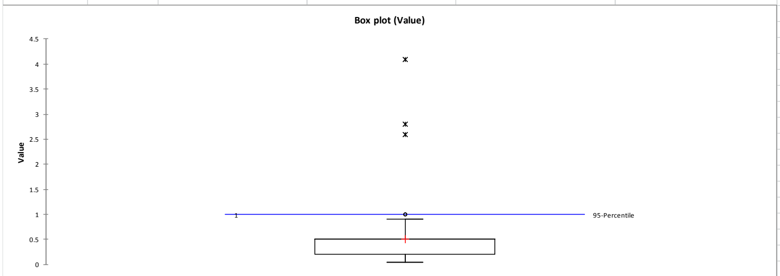
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	4.100				
99%	3.307	2.600	4.100	2.600	4.100
95%	1.000	0.700	4.100	0.700	4.100
90%	0.700	0.500	2.600	0.500	2.800
3rd Quartile 75%	0.500	0.500	0.540	0.500	0.540
Median 50%	0.500	0.200	0.500	0.200	0.500
1st Quartile 25%	0.200	0.100	0.200	0.100	0.200
10%	0.100	0.050	0.140	0.050	0.140
5%	0.050	0.050	0.100	0.050	0.100
1%	0.050	0.050	0.050	0.050	0.100
Minimum 0%	0.050				

Value of the 95-percentile: 1

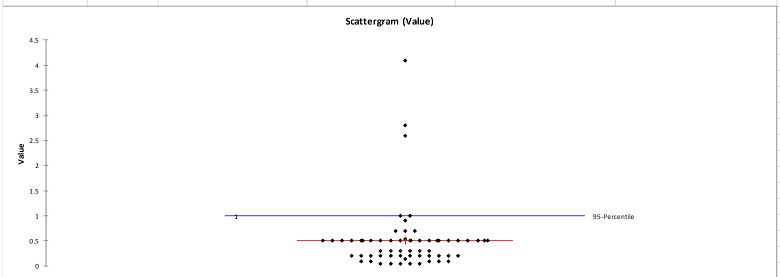
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 11 – Mercury Urban Data

Full spreadsheet for mercury urban original data

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification
1	Mercury	U	1	0.5	48	0.256875	0.219634	0.05	0.07	0.0975	0.175	0.315	0.5	0.5195	1	0.028258602	1.644001683	1.751979592	0.511627907	non-Gaussian
2	Mercury	U	2	0.05												-0.017409482				
3	Mercury	U	3	0.07												-0.012832623				
4	Mercury	U	4	0.14												-0.003139256				
5	Mercury	U	5	0.14												-0.003139256				
6	Mercury	U	6	0.18												-0.000893341				
7	Mercury	U	7	1												0.806951507				
8	Mercury	U	8	1												0.806951507				
9	Mercury	U	9	0.07												-0.012832623				
10	Mercury	U	10	0.07												-0.012832623				
11	Mercury	U	11	0.09												-0.009137674				
12	Mercury	U	12	0.06												-0.015004915				
13	Mercury	U	13	0.06												-0.015004915				
14	Mercury	U	14	0.09												-0.009137674				
15	Mercury	U	15	0.48												0.02184271				
16	Mercury	U	16	0.53												0.040063251				
17	Mercury	U	17	0.4												0.00576512				
24	Mercury	U	18	0.1												-0.00759142				
25	Mercury	U	19	0.08												-0.010880809				
26	Mercury	U	20	0.22												-9.85957E-05				
27	Mercury	U	21	0.36												0.002156521				
28	Mercury	U	22	0.13												-0.004015966				
29	Mercury	U	23	0.5												0.028258602				
30	Mercury	U	24	0.5												0.028258602				
31	Mercury	U	25	0.5												0.028258602				
34	Mercury	U	26	0.05												-0.017409482				
47	Mercury	U	27	0.5												0.028258602				
48	Mercury	U	28	0.5												0.028258602				
50	Mercury	U	29	0.3												0.000157706				
51	Mercury	U	30	0.3												0.000157706				
52	Mercury	U	31	0.3												0.000157706				
53	Mercury	U	32	0.12												-0.005042364				
54	Mercury	U	33	0.14												-0.003139256				
55	Mercury	U	34	0.09												-0.009137674				
56	Mercury	U	35	0.11												-0.00623025				
57	Mercury	U	36	0.14												-0.003139256				
58	Mercury	U	37	0.12												-0.005042364				
59	Mercury	U	38	0.11												-0.00623025				
60	Mercury	U	39	0.06												-0.015004915				
61	Mercury	U	40	0.17												-0.001289277				
62	Mercury	U	41	0.18												-0.000893341				
63	Mercury	U	42	0.18												-0.000893341				
105	Mercury	U	43	0.14												-0.003139256				
107	Mercury	U	44	0.3												0.000157706				
108	Mercury	U	45	0.3												0.000157706				
109	Mercury	U	46	0.3												0.000157706				
110	Mercury	U	47	0.3												0.000157706				
119	Mercury	U	48	0.3												0.000157706				

Full spreadsheet for the log transform data for mercury urban

Log Value	No. of data points	Log Mean	Log Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
-0.301029996	48	-0.728487592	0.351069129	-1.30103	-1.15490196	-1.011439373	-0.75714	-0.503083434	-0.301029996	-0.284581183	0	0.037606244	0.16660866	0.177551504	0.068331811	Symmetrical	-0.3839	-0.18526
-1.301029996												-0.09036574						
-1.15490196												-0.037331576						
-0.853871964												-0.000949098						
-0.853871964												-0.000949098						
-0.744727495												-2.0622E-06						
0												0.18614326						
0												0.18614326						
-1.15490196												-0.037331576						
-1.15490196												-0.037331576						
-1.045757491												-0.015376852						
-1.22184875												-0.057819682						
-1.22184875												-0.057819682						
-1.045757491												-0.015376852						
-0.318758763												0.033118481						
-0.27572413												0.044688415						
-0.397940009												0.017389328						
-1												-0.009637178						
-1.096910013												-0.024077921						
-0.657577319												0.000171675						
-0.443697499												0.011121298						
-0.886056648												-0.001883621						
-0.301029996												0.037606244						
-0.301029996												0.037606244						
-0.301029996												0.037606244						
-1.301029996												-0.09036574						
-0.301029996												0.037606244						
-0.301029996												0.037606244						
-0.522878745												0.004185104						
-0.522878745												0.004185104						
-0.522878745												0.004185104						
-0.920818754												-0.003425546						
-0.853871964												-0.000949098						
-1.045757491												-0.015376852						
-0.958607315												-0.005867355						
-0.853871964												-0.000949098						
-0.920818754												-0.003425546						
-0.958607315												-0.005867355						
-1.22184875												-0.057819682						
-0.769551079												-3.33387E-05						
-0.744727495												-2.0622E-06						
-0.744727495												-2.0622E-06						
-0.853871964												-0.000949098						
-0.522878745												0.004185104						
-0.522878745												0.004185104						
-0.522878745												0.004185104						
-0.522878745												0.004185104						
-0.522878745												0.004185104						

Mercury urban robust percentiles calculation on log transform data

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
48	-0.30103	-0.75714	-0.456109291	-1.375281333	-0.02865	-0.814442676	-0.699835898	-0.70794	-0.69173
	-1.30103		0.543890709						
	-1.1549		0.397762673						
	-0.85387		0.096732678						
	-0.85387		0.096732678						
	-0.74473		-0.012411792						
	0		-0.757139287						
	0		-0.757139287						
	-1.1549		0.397762673						
	-1.1549		0.397762673						
	-1.04576		0.288618204						
	-1.22185		0.464709463						
	-1.22185		0.464709463						
	-1.04576		0.288618204						
	-0.31876		-0.438380524						
	-0.27572		-0.481415156						
	-0.39794		-0.359199278						
	-1		0.242860713						
	-1.09691		0.339770726						
	-0.65758		-0.099561968						
	-0.4437		-0.313441788						
	-0.88606		0.128917361						
	-0.30103		-0.456109291						
	-0.30103		-0.456109291						
	-0.30103		-0.456109291						
	-1.30103		0.543890709						
	-0.30103		-0.456109291						
	-0.30103		-0.456109291						
	-0.52288		-0.234260541						
	-0.52288		-0.234260541						
	-0.52288		-0.234260541						
	-0.92082		0.163679467						
	-0.85387		0.096732678						
	-1.04576		0.288618204						
	-0.95861		0.201468028						
	-0.85387		0.096732678						
	-0.92082		0.163679467						
	-0.95861		0.201468028						
	-1.22185		0.464709463						
	-0.76955		0.012411792						
	-0.74473		-0.012411792						
	-0.74473		-0.012411792						
	-0.85387		0.096732678						
	-0.52288		-0.234260541						
	-0.52288		-0.234260541						
	-0.52288		-0.234260541						
	-0.52288		-0.234260541						
	-0.52288		-0.234260541						

Mercury urban empirical calculation on log transform data

XLSTAT 2013.4.06 - Quantiles estimation - on 27/08/2013 at 13:33:47
Data: Workbook = Hg Urban Sheet.xlsx / Sheet = Empirical / Range = Empirical!$A$1:$A$49 / 48 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

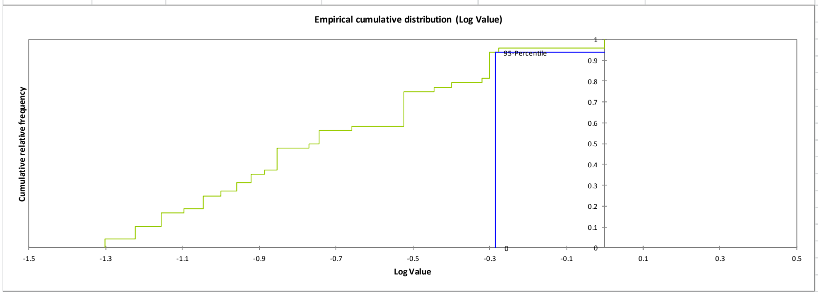
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Log Value	48	0	48	-1.301	0.000	-0.728	0.351

Percentile table (Weighted average at x(Np)):

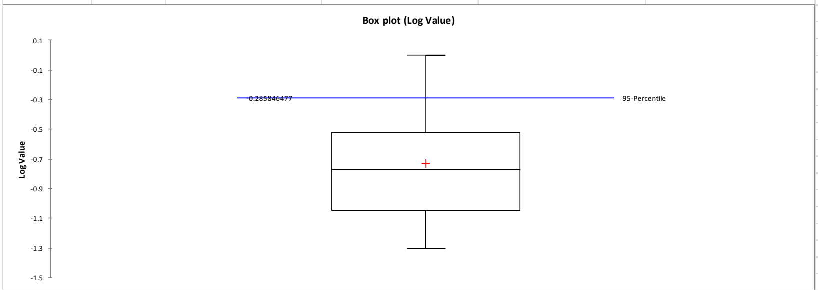
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	0.000				
99%	0.000	-0.276	0.000	-0.276	0.000
95%	-0.286	-0.301	0.000	-0.301	0.000
90%	-0.301	-0.319	0.000	-0.319	0.000
3rd Quartile 75%	-0.523	-0.523	-0.301	-0.523	-0.301
Median 50%	-0.770	-0.921	-0.523	-0.921	-0.523
1st Quartile 25%	-1.046	-1.155	-0.886	-1.155	-0.886
10%	-1.222	-1.301	-1.097	-1.301	-1.155
5%	-1.269	-1.301	-1.222	-1.301	-1.222
1%	-1.301	-1.301	-1.301	-1.301	-1.222
Minimum 0%	-1.301				

Value of the 95-percentile: -0.285846476505119

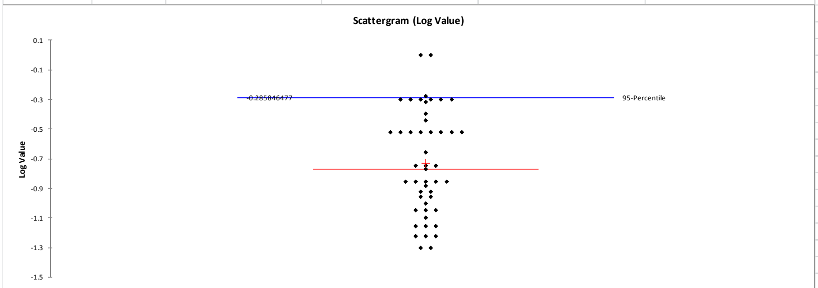
Empirical cumulative distribution:



Box plots:



Scattergrams:



Back transformation of data for mercury urban

Percentile	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	R H
95	-0.28585	-0.30103	0	-0.2845812	-0.3839	-0.18526	-0.69984	-0.70794	-0.69173
Back transformation of data									
Percentile	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	R H
95	0.51779	0.5	1	0.51930059	0.413145	0.652732	0.199602	0.195911	0.203362

Appendix 12 – Nickel Rural Data

Full spreadsheet for nickel rural, including calculation of parametric percentiles

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
18	Nickel	R	1	30	70	18.1228571	3.956619	10.8	14	15	18	20	23	24.55	30	0.386425049	0.326794774	0.34128184	0.111111111	Symmetrical	23.62312	25.47688
19	Nickel	R	2	20												0.001525532						
20	Nickel	R	3	15												-0.007024019						
21	Nickel	R	4	18												-4.27691E-07						
22	Nickel	R	5	20												0.001525532						
23	Nickel	R	6	19												0.000155647						
32	Nickel	R	7	15												-0.007024019						
33	Nickel	R	8	14												-0.016163083						
35	Nickel	R	9	16												-0.002206438						
36	Nickel	R	10	21												0.005493051						
37	Nickel	R	11	20												0.001525532						
38	Nickel	R	12	22												0.013442026						
39	Nickel	R	13	21												0.005493051						
40	Nickel	R	14	23												0.026756282						
41	Nickel	R	15	18												-4.27691E-07						
42	Nickel	R	16	10.8												-0.090567216						
43	Nickel	R	17	13.6												-0.021338699						
44	Nickel	R	18	18												-4.27691E-07						
45	Nickel	R	19	18												-4.27691E-07						
46	Nickel	R	20	20												0.001525532						
64	Nickel	R	21	12												-0.052940948						
65	Nickel	R	22	11												-0.083347395						
66	Nickel	R	23	17												-0.000326515						
67	Nickel	R	24	20												0.001525532						
68	Nickel	R	25	15												-0.007024019						
69	Nickel	R	26	14												-0.016163083						
70	Nickel	R	27	22												0.013442026						
71	Nickel	R	28	20												0.001525532						
72	Nickel	R	29	16												-0.002206438						
73	Nickel	R	30	20												0.001525532						
74	Nickel	R	31	18												-4.27691E-07						
75	Nickel	R	32	15												-0.007024019						
76	Nickel	R	33	20												0.001525532						
77	Nickel	R	34	17												-0.000326515						
78	Nickel	R	35	19												0.000155647						
79	Nickel	R	36	18												-4.27691E-07						
80	Nickel	R	37	17												-0.000326515						
81	Nickel	R	38	17												-0.000326515						
82	Nickel	R	39	14												-0.016163083						
83	Nickel	R	40	16												-0.002206438						
84	Nickel	R	41	21												0.005493051						
85	Nickel	R	42	24												0.04681964						
86	Nickel	R	43	25												0.075015925						
87	Nickel	R	44	23												0.026756282						
88	Nickel	R	45	21												0.005493051						
89	Nickel	R	46	21												0.005493051						
90	Nickel	R	47	18												-4.27691E-07						
91	Nickel	R	48	20												0.001525532						
92	Nickel	R	49	18												-4.27691E-07						
93	Nickel	R	50	23												0.026756282						
94	Nickel	R	51	14												-0.016163083						
95	Nickel	R	52	11												-0.083347395						
96	Nickel	R	53	13												-0.031007451						
97	Nickel	R	54	27												0.161342565						
98	Nickel	R	55	25												0.075015925						
99	Nickel	R	56	20												0.001525532						
100	Nickel	R	57	15												-0.007024019						
101	Nickel	R	58	18												-4.27691E-07						
102	Nickel	R	59	17												-0.000326515						
103	Nickel	R	60	18												-4.27691E-07						
104	Nickel	R	61	23												0.026756282						
106	Nickel	R	62	23.2												0.030184743						
111	Nickel	R	63	18												-4.27691E-07						
112	Nickel	R	64	18												-4.27691E-07						
113	Nickel	R	65	14												-0.016163083						
114	Nickel	R	66	12												-0.052940948						
115	Nickel	R	67	18												-4.27691E-07						
116	Nickel	R	68	14												-0.016163083						
117	Nickel	R	69	12												-0.052940948						
118	Nickel	R	70	14												-0.016163083						

Nickel rural robust percentile calculation

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
70	30	18	-12	-8.6	-0.12286	17.75428571	18.24571429	18.21693	18.27449
	20		-2						
	15		3						
	18		0						
	20		-2						
	19		-1						
	15		3						
	14		4						
	16		2						
	21		-3						
	20		-2						
	22		-4						
	21		-3						
	23		-5						
	18		0						
	10.8		7.2						
	13.6		4.4						
	18		0						
	18		0						
	20		-2						
	12		6						
	11		7						
	17		1						
	20		-2						
	15		3						
	14		4						
	22		-4						
	20		-2						
	16		2						
	20		-2						
	18		0						
	15		3						
	20		-2						
	17		1						
	19		-1						
	18		0						
	17		1						
	17		1						
	14		4						
	16		2						
	21		-3						
	24		-6						
	25		-7						
	23		-5						
	21		-3						
	21		-3						
	18		0						
	20		-2						
	18		0						
	23		-5						
	14		4						
	11		7						
	13		5						
	27		-9						
	25		-7						
	20		-2						
	15		3						
	18		0						
	17		1						
	18		0						
	23		-5						
	23.2		-5.2						
	18		0						
	18		0						
	14		4						
	12		6						
	18		0						
	14		4						
	12		6						
	14		4						

Nickel rural empirical percentile calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 16:37:37
Data: Workbook = Ni Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$71 / 70 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

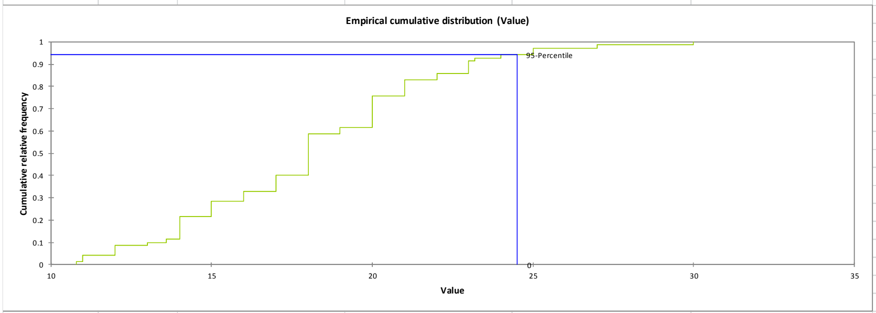
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	70	0	70	10.800	30.000	18.123	3.957

Percentile table (Weighted average at x(Np)):

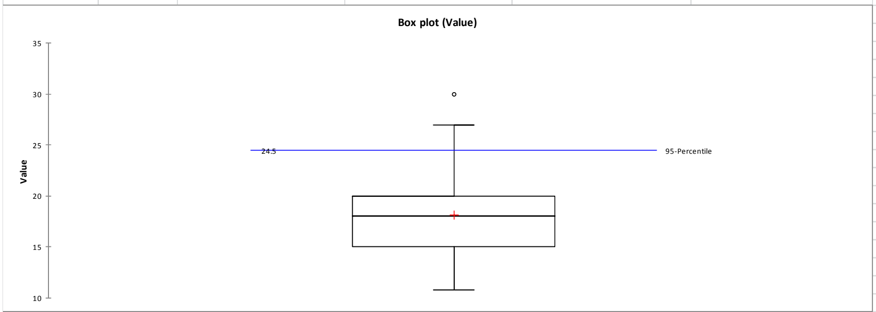
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	30.000				
99%	27.900	25.000	30.000	25.000	30.000
95%	24.500	23.000	30.000	23.000	30.000
90%	23.000	21.000	25.000	21.000	25.000
3rd Quartile 75%	20.000	20.000	22.000	20.000	22.000
Median 50%	18.000	17.000	19.000	17.000	20.000
1st Quartile 25%	15.000	14.000	17.000	14.000	17.000
10%	13.000	11.000	14.000	11.000	14.000
5%	11.500	10.800	13.000	10.800	12.000
1%	10.940	10.800	11.000	10.800	11.000
Minimum 0%	10.800				

Value of the 95-percentile: 24.5

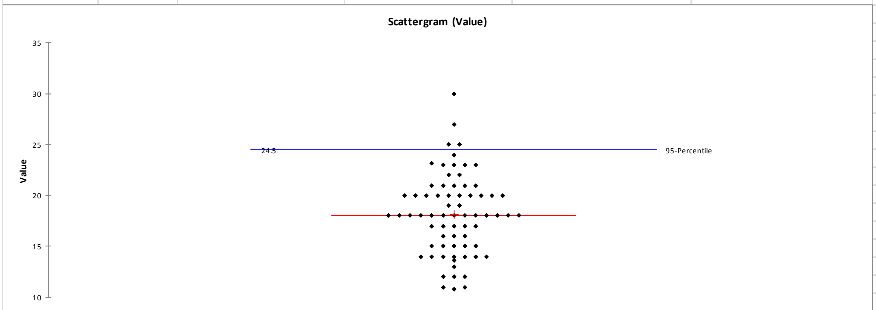
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 13 – Nickel Urban Data

Full spreadsheet for nickel urban, including calculation of parametric percentiles

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
1	Nickel	U	1	23.1	48	23.18541667	9.925038	11	14.875	16.75	22	27.25	32.125	34	72	-1.32798E-08	2.44905231	2.609905884	0.173913043	Gaussian	31.19224	36.80776
2	Nickel	U	2	14												-0.016514249						
3	Nickel	U	3	12												-0.029820762						
4	Nickel	U	4	11												-0.038555238						
5	Nickel	U	5	11												-0.038555238						
6	Nickel	U	6	13												-0.022516383						
7	Nickel	U	7	30												0.006743419						
8	Nickel	U	8	34												0.026952038						
9	Nickel	U	9	16												-0.007905295						
10	Nickel	U	10	16												-0.007905295						
11	Nickel	U	11	16												-0.007905295						
12	Nickel	U	12	16												-0.007905295						
13	Nickel	U	13	17												-0.005042769						
14	Nickel	U	14	19												-0.00156235						
15	Nickel	U	15	13												-0.022516383						
16	Nickel	U	16	18												-0.002971072						
17	Nickel	U	17	17												-0.005042769						
24	Nickel	U	18	21												-0.000222415						
25	Nickel	U	19	20												-0.000688749						
26	Nickel	U	20	22												-3.54957E-05						
27	Nickel	U	21	18												-0.002971072						
28	Nickel	U	22	18												-0.002971072						
29	Nickel	U	23	15												-0.011686504						
30	Nickel	U	24	30												0.006743419						
31	Nickel	U	25	32												0.014593773						
34	Nickel	U	26	26												0.000475121						
47	Nickel	U	27	24												1.15178E-05						
48	Nickel	U	28	24												1.15178E-05						
50	Nickel	U	29	38												0.069283618						
51	Nickel	U	30	15												-0.011686504						
52	Nickel	U	31	31												0.010169034						
53	Nickel	U	32	22												-3.54957E-05						
54	Nickel	U	33	24												1.15178E-05						
55	Nickel	U	34	28												0.002378145						
56	Nickel	U	35	23												-1.35834E-07						
57	Nickel	U	36	23												-1.35834E-07						
58	Nickel	U	37	21												-0.000222415						
59	Nickel	U	38	21												-0.000222415						
60	Nickel	U	39	27												0.001182779						
61	Nickel	U	40	23												-1.35834E-07						
62	Nickel	U	41	24												1.15178E-05						
63	Nickel	U	42	22												-3.54957E-05						
105	Nickel	U	43	21.8												-5.66635E-05						
107	Nickel	U	44	31												0.010169034						
108	Nickel	U	45	34												0.026952038						
109	Nickel	U	46	33												0.02014549						
110	Nickel	U	47	33												0.02014549						
119	Nickel	U	48	72												2.478625954						

Nickel urban robust percentile calculation

No. of values	Value	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
48	23.1	22	-1.1	-56.9	-1.18542	19.62916667	24.37083333	24.03548	24.70618
	14		8						
	12		10						
	11		11						
	11		11						
	13		9						
	30		-8						
	34		-12						
	16		6						
	16		6						
	16		6						
	16		6						
	17		5						
	19		3						
	13		9						
	18		4						
	17		5						
	21		1						
	20		2						
	22		0						
	18		4						
	18		4						
	15		7						
	30		-8						
	32		-10						
	26		-4						
	24		-2						
	24		-2						
	38		-16						
	15		7						
	31		-9						
	22		0						
	24		-2						
	28		-6						
	23		-1						
	23		-1						
	21		1						
	21		1						
	27		-5						
	23		-1						
	24		-2						
	22		0						
	21.8		0.2						
	31		-9						
	34		-12						
	33		-11						
	33		-11						
	72		-50						

Nickel urban empirical percentile calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 27/08/2013 at 13:39:24
Data: Workbook = Ni Urban Sheet.xlsx / Sheet = Empirical / Range = Empirical!$A$1:$A$49 / 48 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

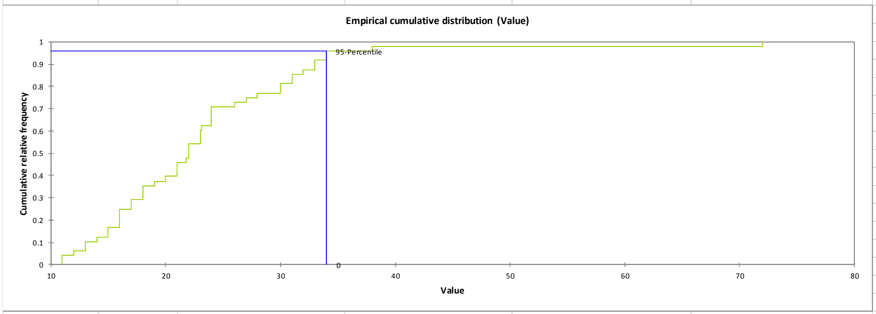
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	48	0	48	11.000	72.000	23.185	9.925

Percentile table (Weighted average at x(Np)):

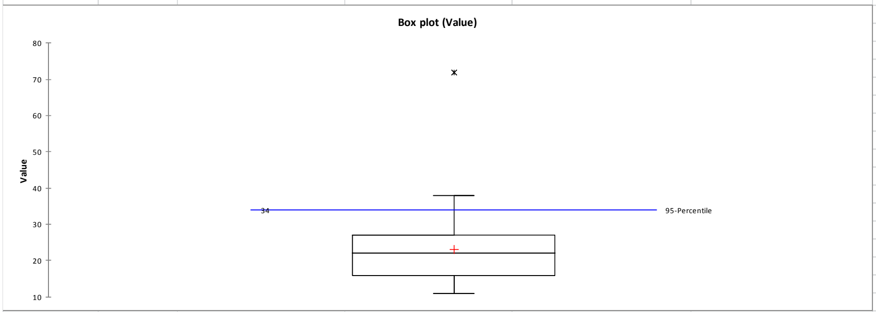
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	72.000				
99%	55.680	34.000	72.000	34.000	72.000
95%	34.000	33.000	72.000	32.000	72.000
90%	33.000	30.000	38.000	30.000	72.000
3rd Quartile 75%	27.000	23.100	32.000	23.100	32.000
Median 50%	22.000	18.000	24.000	18.000	24.000
1st Quartile 25%	16.000	14.000	19.000	14.000	19.000
10%	13.000	11.000	16.000	11.000	15.000
5%	11.400	11.000	13.000	11.000	12.000
1%	11.000	11.000	11.000	11.000	12.000
Minimum 0%	11.000				

Value of the 95-percentile: 34

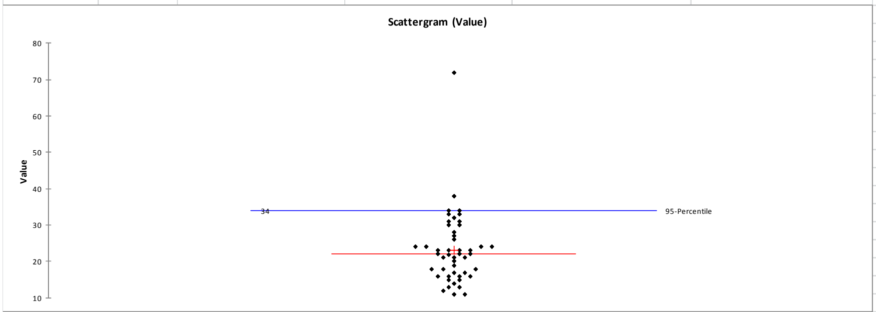
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 14 – Lead Rural Data

Full spreadsheet for lead rural, including parametric percentile calculation on the original data

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
18	Lead	R	1	140	70	140.46	173.101	34	44.625	49	63	140	264.375	585.5	880	-2.68087E-10	2.347152003	2.451203072	0.832764505	non-Gaussian	544.9493	626.0507
19	Lead	R	2	55												-0.001719062						
20	Lead	R	3	45												-0.002395892						
21	Lead	R	4	49												-0.002107155						
22	Lead	R	5	61												-0.001381811						
23	Lead	R	6	64												-0.001231136						
32	Lead	R	7	43												-0.00254966						
33	Lead	R	8	46												-0.002321383						
35	Lead	R	9	148												1.18064E-06						
36	Lead	R	10	246												0.003237832						
37	Lead	R	11	249												0.003521863						
38	Lead	R	12	617												0.2980582						
39	Lead	R	13	386												0.040772726						
40	Lead	R	14	386												0.040772726						
41	Lead	R	15	142												1.00592E-08						
42	Lead	R	16	35.3												-0.003202984						
43	Lead	R	17	40.8												-0.00272625						
44	Lead	R	18	70												-0.000963454						
45	Lead	R	19	72												-0.000883718						
46	Lead	R	20	83												-0.000522516						
64	Lead	R	21	46												-0.002321383						
65	Lead	R	22	39												-0.002876654						
66	Lead	R	23	34												-0.003323245						
67	Lead	R	24	61												-0.001381811						
68	Lead	R	25	45												-0.002395892						
69	Lead	R	26	59												-0.001488799						
70	Lead	R	27	53												-0.001842601						
71	Lead	R	28	49												-0.002107155						
72	Lead	R	29	48												-0.002177031						
73	Lead	R	30	51												-0.001971922						
74	Lead	R	31	64												-0.001231136						
75	Lead	R	32	50												-0.002038791						
76	Lead	R	33	76												-0.00073769						
77	Lead	R	34	51												-0.001971922						
78	Lead	R	35	75												-0.000772558						
79	Lead	R	36	61												-0.001381811						
80	Lead	R	37	46												-0.002321383						
81	Lead	R	38	41												-0.002709869						
82	Lead	R	39	75												-0.000772558						
83	Lead	R	40	150												2.39138E-06						
84	Lead	R	41	230												0.001977217						
85	Lead	R	42	620												0.303722881						
86	Lead	R	43	880												1.114007206						
87	Lead	R	44	500												0.128010218						
88	Lead	R	45	110												-7.78382E-05						
89	Lead	R	46	80												-0.000608706						
90	Lead	R	47	590												0.250211932						
91	Lead	R	48	580												0.233882757						
92	Lead	R	49	140												-2.68087E-10						
93	Lead	R	50	220												0.001385989						
94	Lead	R	51	44												-0.002471979						
95	Lead	R	52	49												-0.002107155						
96	Lead	R	53	59												-0.001488799						
97	Lead	R	54	44												-0.002471979						
98	Lead	R	55	47												-0.002248435						
99	Lead	R	56	47												-0.002248435						
100	Lead	R	57	58												-0.001544305						
101	Lead	R	58	60												-0.00143464						
102	Lead	R	59	68												-0.001047847						
103	Lead	R	60	60												-0.00143464						
104	Lead	R	61	240												0.002716414						
106	Lead	R	62	59.1												-0.001483323						
111	Lead	R	63	72												-0.000883718						
112	Lead	R	64	79												-0.000639412						
113	Lead	R	65	62												-0.001330295						
114	Lead	R	66	42												-0.002628951						
115	Lead	R	67	290												0.009210329						
116	Lead	R	68	90												-0.000353871						
117	Lead	R	69	130												-3.15209E-06						
118	Lead	R	70	130												-3.15209E-06						

Full spreadsheet for lead rural log transform data

Log Value	No. of data points	Log Mean	Log Std Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification
2.146128036	70	1.954650314	0.364624087	1.53147892	1.649552575	1.69019608	1.799286	2.146128036	2.421023841	2.767511204	2.9444827	0.00206881	1.148008618	1.198900731	0.611824151	non-Gaussian
1.740362689												-0.002899727				
1.653212514												-0.008071589				
1.69019608												-0.005450263				
1.785329835												-0.001430521				
1.806179974												-0.000964462				
1.633468456												-0.009763801				
1.662757832												-0.00732883				
2.170261715												0.0029538				
2.390935107												0.024472375				
2.396199347												0.025368963				
2.790285164												0.171955285				
2.586587305												0.074368193				
2.586587305												0.074368193				
2.152288344												0.002274979				
1.547774705												-0.019849559				
1.610660163												-0.011995101				
1.84509804												-0.000387463				
1.857332496												-0.000271608				
1.919078092												-1.32647E-05				
1.662757832												-0.00732883				
1.591064607												-0.014164013				
1.531478917												-0.022331341				
1.785329835												-0.001430521				
1.653212514												-0.008071589				
1.770852012												-0.001829744				
1.72427587												-0.003603038				
1.69019608												-0.005450263				
1.681241237												-0.006022887				
1.707570176												-0.004445078				
1.806179974												-0.000964462				
1.698970004												-0.004925584				
1.880813592												-0.000118627				
1.707570176												-0.004445078				
1.875061263												-0.000148568				
1.785329835												-0.001430521				
1.662757832												-0.00732883				
1.612783857												-0.011774307				
1.875061263												-0.000148568				
2.176091259												0.003199924				
2.361727836												0.019879125				
2.792391689												0.173258995				
2.944482672												0.285792506				
2.698970004												0.121519065				
2.041392685												0.000192336				
1.903089987												-4.03936E-05				
2.770852012												0.160235366				
2.763427994												0.155902603				
2.146128036												0.00206881				
2.342422681												0.0171829				
1.643452676												-0.008881264				
1.69019608												-0.005450263				
1.770852012												-0.001829744				
1.643452676												-0.008881264				
1.672097858												-0.006647574				
1.672097858												-0.006647574				
1.763427994												-0.002060543				
1.77815125												-0.001620291				
1.832508913												-0.000536976				
1.77815125												-0.001620291				
2.380211242												0.022711778				
1.771587481												-0.001807867				
1.857332496												-0.000271608				
1.897627091												-5.46413E-05				
1.792391689												-0.001258894				
1.62324929												-0.010725744				
2.462397998												0.038575339				
1.954242509												-1.99859E-11				
2.113943352												0.001191122				
2.113943352												0.001191122				

Full spreadsheet for lead rural Box Cox transform data

Box-Cox Value	No. of data points	Box-Cox Mean	Box-Cox Std. Dev.	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification
11.83215957	70	10.49685808	5.542089186	5.83095189	6.680096051	7	7.937004	11.83215957	16.24835354	24.1968887	29.664794	0.000199811	1.723570486	1.799977703	0.737275489	non-Gaussian
7.416198487												-0.00245365				
6.708203932												-0.004563902				
7												-0.003588521				
7.810249676												-0.001627401				
8												-0.00130636				
6.557438524												-0.005130718				
6.782329983												-0.004301227				
12.16552506												0.000389933				
15.68438714												0.01171556				
15.77973384												0.012373502				
24.8394847												0.247609616				
19.6468827												0.064290709				
19.6468827												0.064290709				
11.91637529												0.000240051				
5.941380311												-0.007933836				
6.387487769												-0.005823811				
8.366600265												-0.000811293				
8.485281374												-0.000683111				
9.110433579												-0.000223651				
6.782329983												-0.004301227				
6.244997998												-0.006450871				
5.830951895												-0.008524903				
7.810249676												-0.001627401				
6.708203932												-0.004563902				
7.681145748												-0.001873468				
7.280109889												-0.002793397				
7												-0.003588521				
6.92820323												-0.003814126				
7.141428429												-0.003170486				
8												-0.00130636				
7.071067812												-0.003374145				
8.717797887												-0.000472556				
7.141428429												-0.003170486				
8.660254038												-0.00051991				
7.810249676												-0.001627401				
6.782329983												-0.004301227				
6.403124237												-0.005757583				
8.660254038												-0.00051991				
12.24744871												0.000450231				
15.16575089												0.008541284				
24.8997992												0.25074656				
29.66479395												0.591027205				
22.36067977												0.140137789				
10.48808848												-5.66006E-11				
8.94427191												-0.000314086				
24.2899156												0.220223201				
24.08318916												0.210468938				
11.83215957												0.000199811				
14.83239697												0.006839285				
6.633249581												-0.004840172				
7												-0.003588521				
7.681145748												-0.001873468				
6.633249581												-0.004840172				
6.8556546												-0.004051504				
6.8556546												-0.004051504				
7.615773106												-0.00200701				
7.745966692												-0.001747036				
8.246211251												-0.000956761				
7.745966692												-0.001747036				
15.49193338												0.010459419				
7.687652438												-0.00186051				
8.485281374												-0.000683111				
8.888194417												-0.000349363				
7.874007874												-0.001514265				
6.480740698												-0.005436265				
17.02938637												0.023395123				
9.486832981												-8.64725E-05				
11.40175425												6.21839E-05				
11.40175425												6.21839E-05				

Lead rural robust percentile calculation on original data

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
70	140	63	-77	-5422.2	-77.46	-91.92	217.92	199.7742	236.0658
	55		8						
	45		18						
	49		14						
	61		2						
	64		-1						
	43		20						
	46		17						
	148		-85						
	246		-183						
	249		-186						
	617		-554						
	386		-323						
	386		-323						
	142		-79						
	35.3		27.7						
	40.8		22.2						
	70		-7						
	72		-9						
	83		-20						
	46		17						
	39		24						
	34		29						
	61		2						
	45		18						
	59		4						
	53		10						
	49		14						
	48		15						
	51		12						
	64		-1						
	50		13						
	76		-13						
	51		12						
	75		-12						
	61		2						
	46		17						
	41		22						
	75		-12						
	150		-87						
	230		-167						
	620		-557						
	880		-817						
	500		-437						
	110		-47						
	80		-17						
	590		-527						
	580		-517						
	140		-77						
	220		-157						
	44		19						
	49		14						
	59		4						
	44		19						
	47		16						
	47		16						
	58		5						
	60		3						
	68		-5						
	60		3						
	240		-177						
	59.1		3.9						
	72		-9						
	79		-16						
	62		1						
	42		21						
	290		-227						
	90		-27						
	130		-67						
	130		-67						

Lead rural empirical percentile calculation on original data

XLSTAT 2013.4.06 - Quantiles estimation - on 22/08/2013 at 16:42:29
Data: Workbook = Pb Rural Sheet.xlsx / Sheet = Sheet1 / Range = Sheet1!$A$1:$A$71 / 70 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

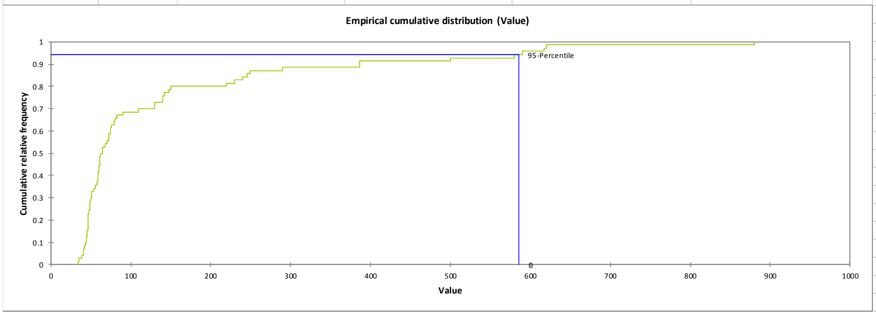
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Value	70	0	70	34.000	880.000	140.460	173.101

Percentile table (Weighted average at x(Np)):

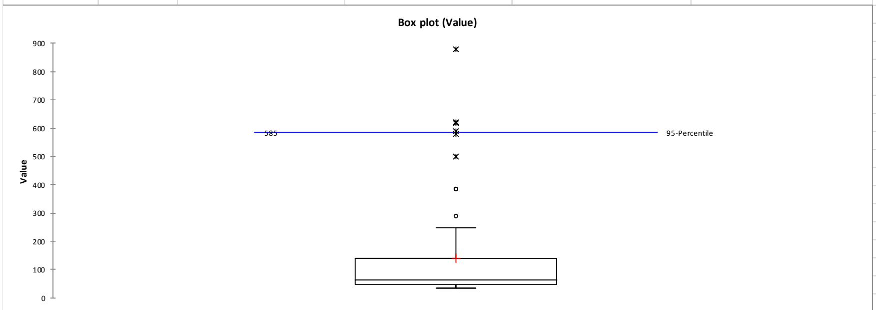
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	880.000				
99%	698.000	617.000	880.000	617.000	880.000
95%	585.000	386.000	880.000	290.000	880.000
90%	386.000	230.000	617.000	230.000	617.000
3rd Quartile 75%	140.000	79.000	246.000	79.000	246.000
Median 50%	62.000	59.000	75.000	59.000	76.000
1st Quartile 25%	48.500	45.000	55.000	45.000	55.000
10%	43.000	35.300	46.000	35.300	46.000
5%	39.900	34.000	43.000	34.000	41.000
1%	34.910	34.000	35.300	34.000	35.300
Minimum 0%	34.000				

Value of the 95-percentile: 585

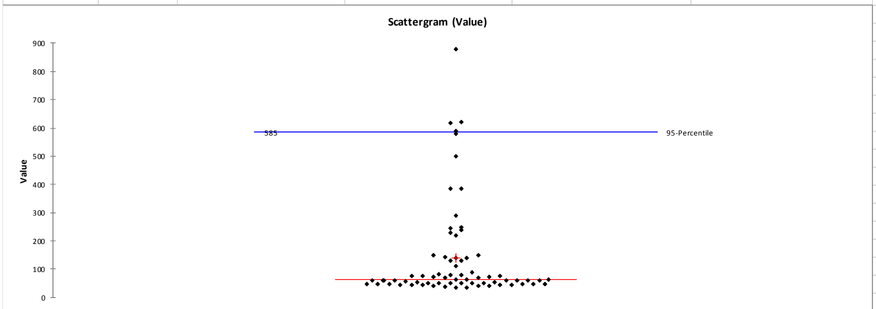
Empirical cumulative distribution:



Box plots:



Scattergrams:



Appendix 15 – Lead Rural Data

Full spreadsheet for lead urban original data

ID	Contaminant	Area	Data Point No.	Value	No. of data points	Mean	Std. Dev	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-μ)^3/Nσ^3	SC = SUM(xi-μ)^3/Nσ^3	Skew function value	Octile Skew	Classification
1	Lead	U	1	211.5	48	77.39375	66.57771	1	14	35.75	62	90	172.5	217.025	260	0.170262637	1.164610229	1.241101742	0.394321767	non-Gaussian
2	Lead	U	2	43												-0.002872175				
3	Lead	U	3	45												-0.002399695				
4	Lead	U	4	43												-0.002872175				
5	Lead	U	5	190												0.100799827				
6	Lead	U	6	65												-0.000134394				
7	Lead	U	7	32												-0.006603292				
8	Lead	U	8	28												-0.008507232				
9	Lead	U	9	40												-0.003691215				
10	Lead	U	10	38												-0.00431573				
11	Lead	U	11	52												-0.001155987				
12	Lead	U	12	38												-0.00431573				
13	Lead	U	13	43												-0.002872175				
14	Lead	U	14	37												-0.004652805				
15	Lead	U	15	260												0.429852259				
16	Lead	U	16	240												0.303517686				
17	Lead	U	17	170												0.056065177				
24	Lead	U	18	220												0.204733061				
25	Lead	U	19	200												0.130109836				
26	Lead	U	20	140												0.01732309				
27	Lead	U	21	88												8.42283E-05				
28	Lead	U	22	130												0.010277425				
29	Lead	U	23	62												-0.000257517				
30	Lead	U	24	62												-0.000257517				
31	Lead	U	25	24												-0.010745919				
34	Lead	U	26	18												-0.014790898				
47	Lead	U	27	133												0.012137888				
48	Lead	U	28	93												0.000268329				
50	Lead	U	30	7												-0.024624907				
51	Lead	U	31	7												-0.024624907				
52	Lead	U	32	30												-0.007515113				
53	Lead	U	33	75												-9.68297E-07				
54	Lead	U	34	85												3.10659E-05				
55	Lead	U	35	89												0.000110369				
56	Lead	U	36	80												1.24974E-06				
57	Lead	U	37	95												0.000385277				
58	Lead	U	38	78												1.573E-08				
59	Lead	U	39	61												-0.000311034				
60	Lead	U	40	58												-0.000514941				
61	Lead	U	41	67												-7.92664E-05				
62	Lead	U	42	83												1.24391E-05				
63	Lead	U	43	72												-1.10776E-05				
105	Lead	U	44	63.4												-0.000193453				
107	Lead	U	45	1												-0.031473554				
108	Lead	U	46	1												-0.031473554				
109	Lead	U	47	1												-0.031473554				
110	Lead	U	48	1												-0.031473554				
119	Lead	U	49	15												-0.017147293				

Full spreadsheet for lead urban log transformed data including parametric percentile calculation

Log Value	No. of data points	Log Mean	Log Std. Dev.	Minimum	0.125 percentile	25th percentile	Median	75th percentile	0.875 percentile	95th percentile	Maximum	(xi-µ)^3/Nσ^3	SC = SUM(xi-µ)^3/Nσ^3	Skew function value	Octile Skew	Classification	LCL	UCL
2.325310372	48	1.645385352	0.611141635	0	1.134717107	1.552438788	1.792392	1.954163242	2.236487006	2.336433373	2.4149733	0.028689053	-1.479807094	-1.577000715	-0.193851063	Symmetrical	2.163544	2.509323
1.633468456												-1.54462E-07						
1.653212514												4.37669E-08						
1.633468456												-1.54462E-07						
2.278753601												0.023190067						
1.812913357												0.000429136						
1.505149978												-0.000251713						
1.447158031												-0.000710925						
1.602059991												-7.42266E-06						
1.579783597												-2.57679E-05						
1.716003344												3.21425E-05						
1.579783597												-2.57679E-05						
1.633468456												-1.54462E-07						
1.568201724												-4.19671E-05						
2.414973348												0.041601405						
2.380211242												0.036214834						
2.230448921												0.018278574						
2.342422681												0.03091016						
2.301029996												0.025724012						
2.146128036												0.011459799						
1.944482672												0.002442141						
2.113943352												0.009389084						
1.792391689												0.000289962						
1.792391689												0.000289962						
1.380211242												-0.00170187						
1.255272505												-0.00541881						
2.123851641												0.009997403						
1.968482949												0.003078466						
0.84509804												-0.046781155						
0.84509804												-0.046781155						
1.477121255												-0.000434818						
1.875061263												0.001105807						
1.929418926												0.002091424						
1.949390007												0.00256433						
1.903089987												0.001562068						
1.977723605												0.003350226						
1.892094603												0.001370533						
1.785329835												0.00025015						
1.763427994												0.000150124						
1.826074803												0.000538433						
1.919078092												0.001871212						
1.857332496												0.000868992						
1.802089258												0.000351215						
0												-0.406570653						
0												-0.406570653						
0												-0.406570653						
0												-0.406570653						
1.176091259												-0.009433403						

Lead urban robust percentile calculation on log transformed data

No. of values	Values	Median	Median-value	SUM(median-value)	MAD	95th percentile = Median+2MAD	95th percentile = Median-2MAD	LCL	UCL
48	2.32531	1.792392	-0.532918682	7.056304199	0.147006	2.086404364	1.498379015	2.044817	2.127992
	1.633468		0.158923234						
	1.653213		0.139179176						
	1.633468		0.158923234						
	2.278754		-0.486361911						
	1.812913		-0.020521667						
	1.50515		0.287241711						
	1.447158		0.345233658						
	1.60206		0.190331698						
	1.579784		0.212608093						
	1.716003		0.076388346						
	1.579784		0.212608093						
	1.633468		0.158923234						
	1.568202		0.224189965						
	2.414973		-0.622581658						
	2.380211		-0.587819552						
	2.230449		-0.438057232						
	2.342423		-0.550030991						
	2.30103		-0.508638306						
	2.146128		-0.353736346						
	1.944483		-0.152090983						
	2.113943		-0.321551663						
	1.792392		0						
	1.792392		0						
	1.380211		0.412180448						
	1.255273		0.537119184						
	2.123852		-0.331459951						
	1.968483		-0.176091259						
	0.845098		0.947293649						
	0.845098		0.947293649						
	1.477121		0.315270435						
	1.875061		-0.082669574						
	1.929419		-0.137027236						
	1.94939		-0.156998317						
	1.90309		-0.110698297						
	1.977724		-0.185331916						
	1.892095		-0.099702913						
	1.78533		0.007061854						
	1.763428		0.028963696						
	1.826075		-0.033683113						
	1.919078		-0.126686403						
	1.857332		-0.064940807						
	1.802089		-0.009697568						
	0		1.792391689						
	0		1.792391689						
	0		1.792391689						
	0		1.792391689						
	1.176091		0.61630043						

Lead urban empirical percentile calculation

XLSTAT 2013.4.06 - Quantiles estimation - on 27/08/2013 at 13:46:36
Data: Workbook = Pb Urban Sheet.xlsx / Sheet = Empirical / Range = Empirical!$A$1:$A$49 / 48 rows and 1 column
Significance level (%): 5
Percentile: 95

Summary statistics:

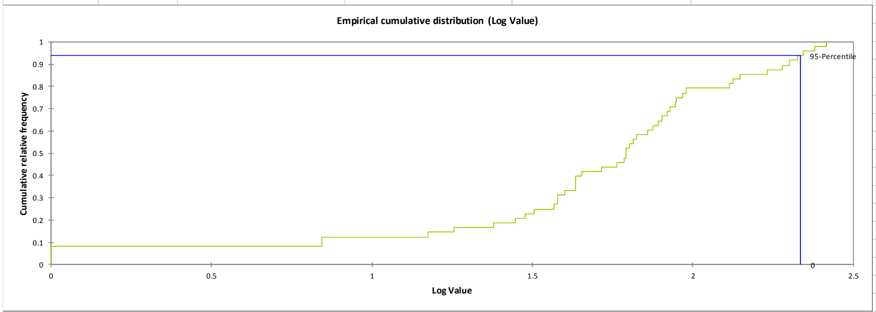
Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Log Value	48	0	48	0.000	2.415	1.645	0.611

Percentile table (Weighted average at x(Np)):

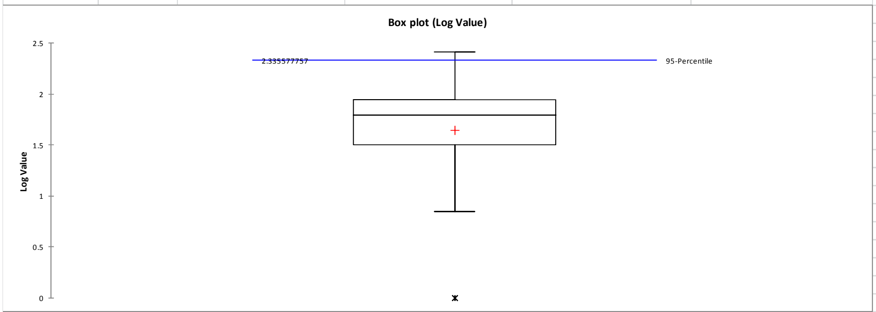
Percentile	Value	Lower bound (Normal based)	Upper bound (Normal based)	Lower bound (Distribution free)	Upper bound (Distribution free)
Maximum 100%	2.415				
99%	2.398	2.342	2.415	2.342	2.415
95%	2.336	2.279	2.415	2.230	2.415
90%	2.283	2.114	2.380	2.114	2.415
3rd Quartile 75%	1.949	1.875	2.230	1.875	2.230
Median 50%	1.792	1.633	1.892	1.633	1.892
1st Quartile 25%	1.505	0.845	1.633	0.845	1.633
10%	0.676	0.000	1.380	0.000	1.176
5%	0.000	0.000	0.845	0.000	0.845
1%	0.000	0.000	0.000	0.000	0.845
Minimum 0%	0.000				

Value of the 95-percentile: 2.33557775717775

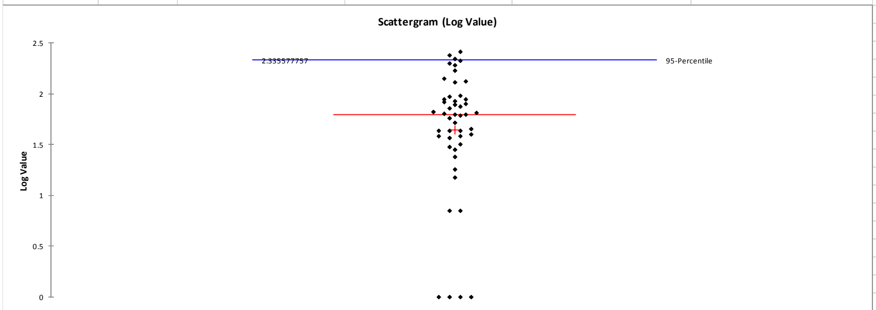
Empirical cumulative distribution:



Box plots:



Scattergrams:



Lead urban back transformation of log transformed data

Percentile	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	R H
95	2.335578	2.278754	2.414973	2.33643337	2.163544	2.509323	2.086404	2.044817	2.127992
Back transformation of data									
Percentile	Empirical	Emp L	Emp H	Parametric	P L	P H	Robust	R L	R H
95	216.5598	190	260	216.986829	145.7282	323.0897	122.0125	110.8707	134.274

DARLINGTON BOROUGH COUNCIL
Environmental Health, Town Hall, Darlington, DL1 5QT
Tel: 01325 388563 Fax: 01325 388446
Email: environmentalhealth@darlington.gov.uk
Website: www.darlington.gov.uk