Statistical Analysis of Beryllium in Settled Dust

Joint EFCOG/DOE Chemical Management Workshop

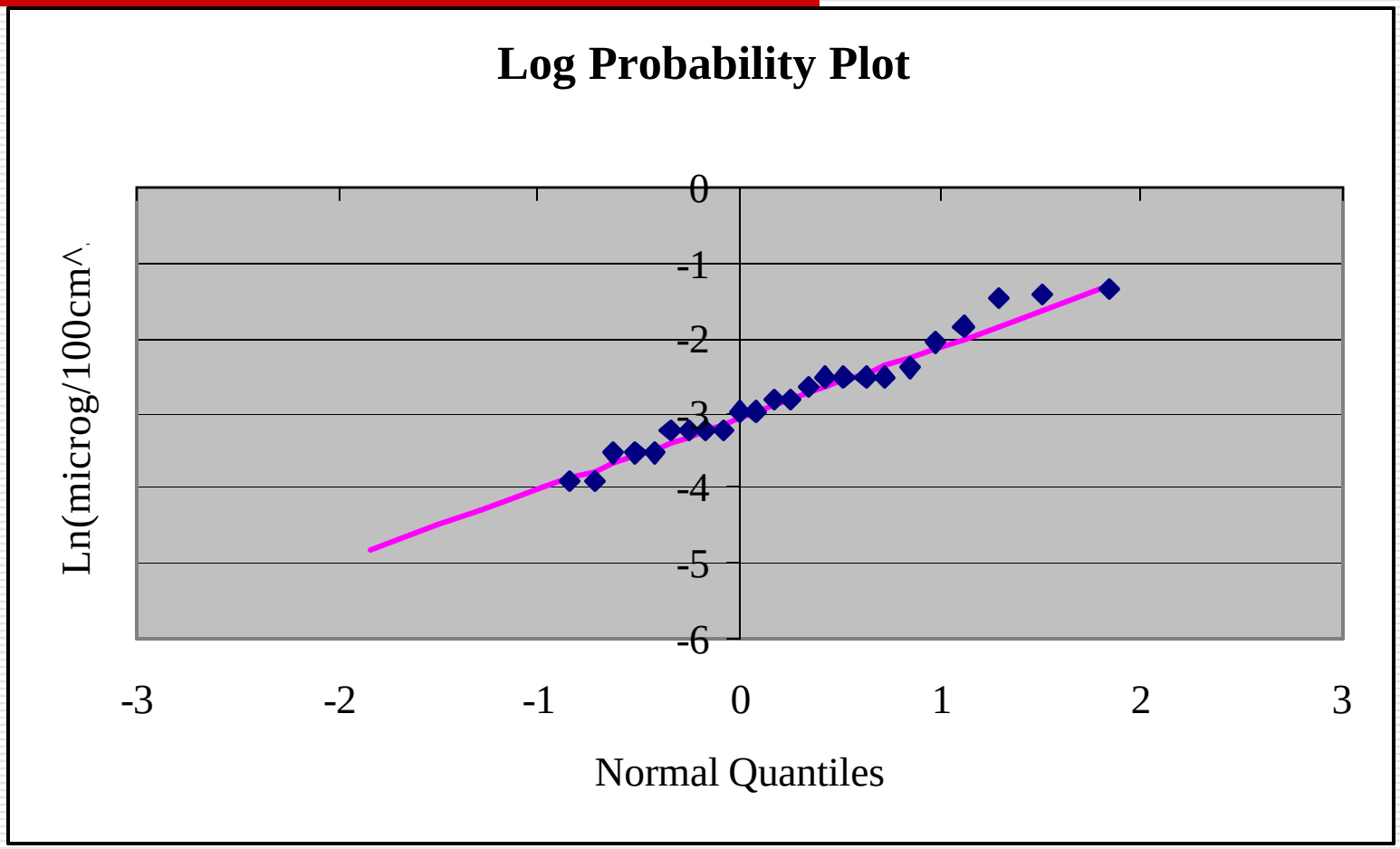
March 14–16, 2006

Surface Loading - µg/100cm^2

- Thought to be a risk indicator
  - Beryllium available for re-entrainment
- Results are compared to a limit of 0.2 µg/100cm^2
  - Clean surfaces should have beryllium loading less than this.
  - Z-test, are results outside the normal range
- Laboratory reporting limits are about 0.03 µg/100cm^2
  - Samples from clean surfaces are mostly non-detects

Surface Loading Data

R^2 = 0.97 Non-detects = 17%



Tolerance Limits

- Parametric - Maximum Likelihood Estimate
  - The geometric 95-95 upper tolerance limit
  - 95-95 UTL = 0.37 µg/100cm^2
- Non-Parametric – Order Statistic
  - Largest of 59, second largest of 93, etc.
  - N = 29, Distribution-free estimate of the 95-95 UTL not available

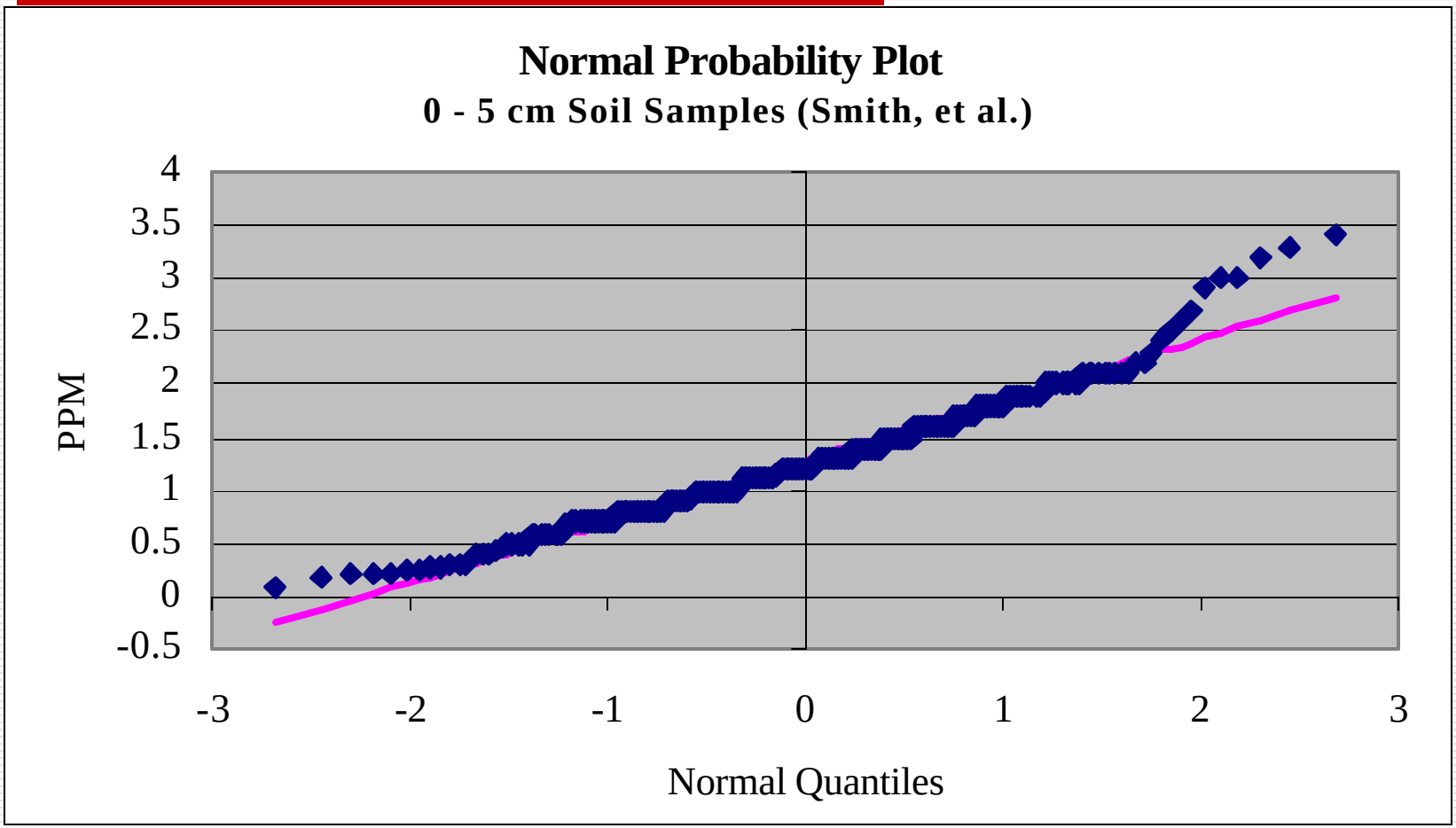
Exceedance Tests

- Parametric - The MLE of the percent of values exceeding the specified limit = 6.4%
  - 95% lower confidence limit = 2.2%
  - 95% upper confidence limit = 15%
- Non-parametric - Binomial estimate of the percent of values exceeding the limit = 10.3%
  - 95% lower confidence limit = 2.9%
  - 95% upper confidence limit = 24.6%

Concentration - ppm (µg/g)

- Thought to be an indicator of cause
  - Concentration higher near the source
- Is concentration in settled dust different from nearby soils
  - Compare settled dust to reference area
  - T-test, are samples from the same population
- Laboratory reporting limit about 0.1 ppm
  - Bulk settled dust has fewer non-detect results

Concentrations in Soils
USGS Open-File Report 2005–1253



Mean and Confidence Interval

- For complete data – arithmetic mean and confidence interval
  - Upper 95% confidence limit – lower 95% confidence limit = 90% confidence interval
- For data with non-detects Kaplan-Meier mean and confidence interval

Beryllium Concentration in Soil USGS Open-File Report 2005–1253

- 275 soil sample collected along East-West and North-South Continental Transects

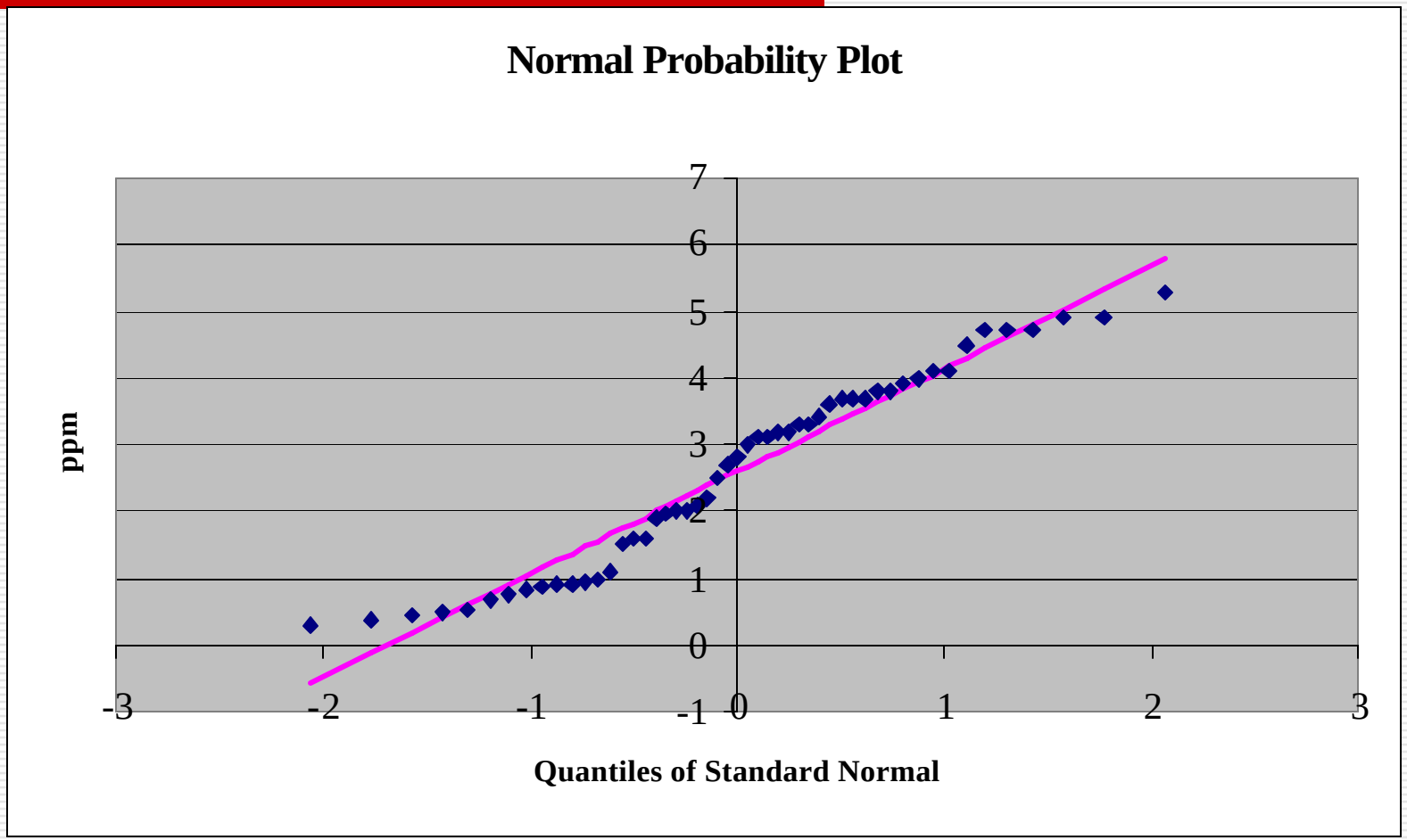
- Arithmetic Mean 1.28 ppm

- 90% CI: 1.22 - 1.34 ppm

- 95th Percentile 2.23 ppm

Be Concentration in Carpet Dust

R^2 = 0.96 Non-detects = 0%

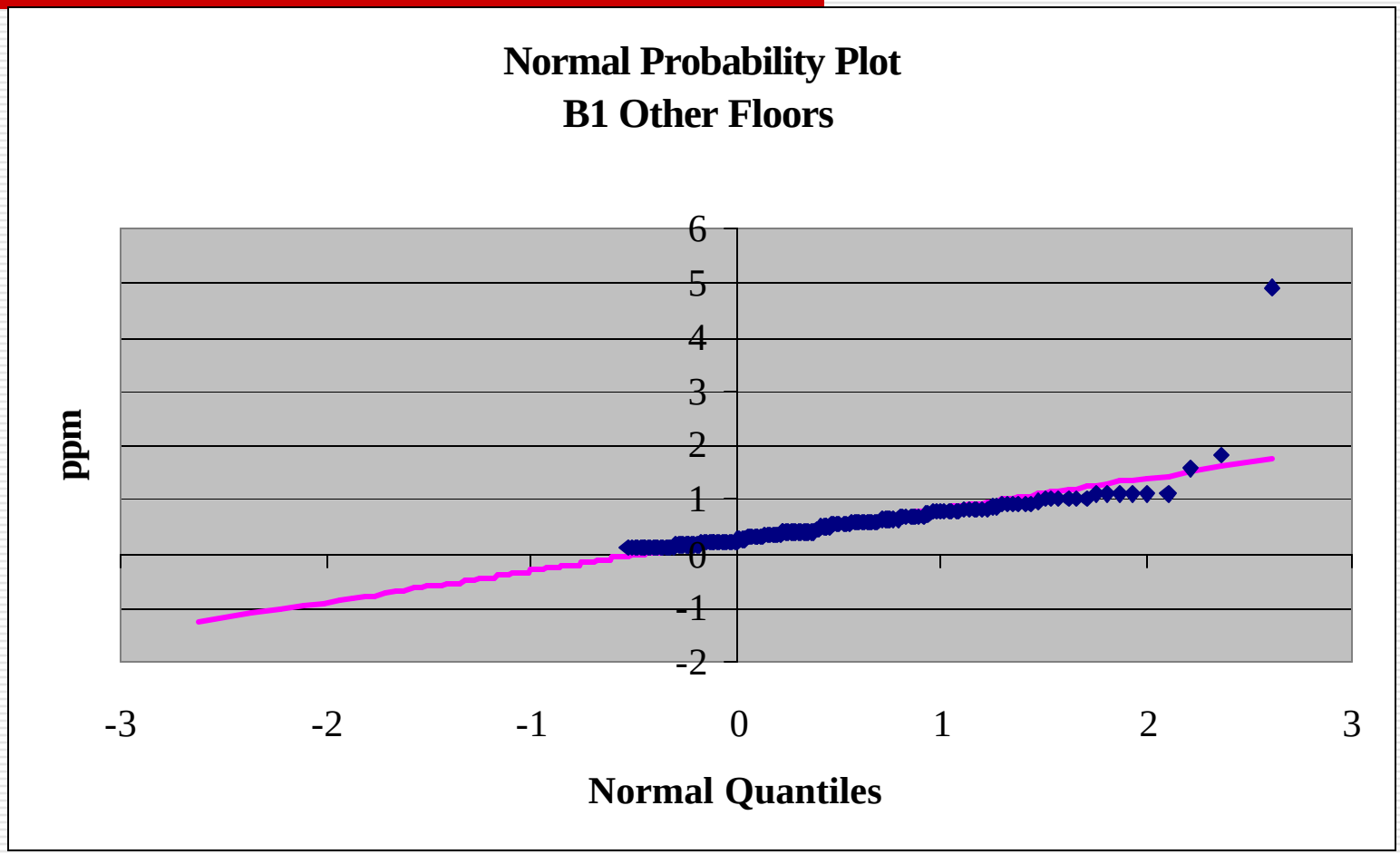


Be Concentration in Carpet Dust B1 First Floor

- Arithmetic Mean 2.60 ppm
- 90% CI: 2.41 – 2.79 ppm
- 90% CI: US Soil 1.22 - 1.34 ppm
- Confidence intervals do not overlap
  1st Floor significantly larger then soil

Be Concentration in Carpet Dust

R^2 = 0.69 Non-detects = 30%



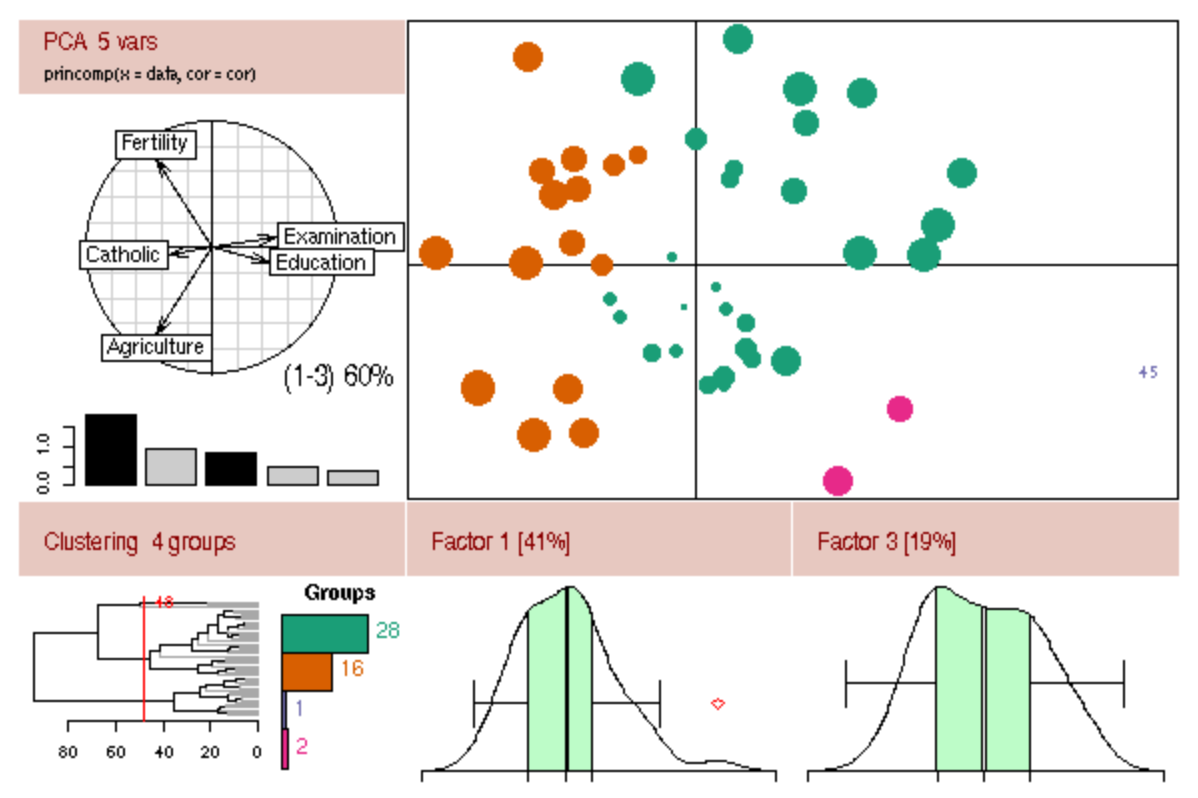
Be Concentration in Carpet Dust B1 Other Floors

- Arithmetic Mean 0.37 ppm
- 90% CI: 0.32 – 0.42 ppm
- 90% CI: US Soil 1.22 - 1.34 ppm
- Confidence intervals do not overlap Other Floors significantly smaller

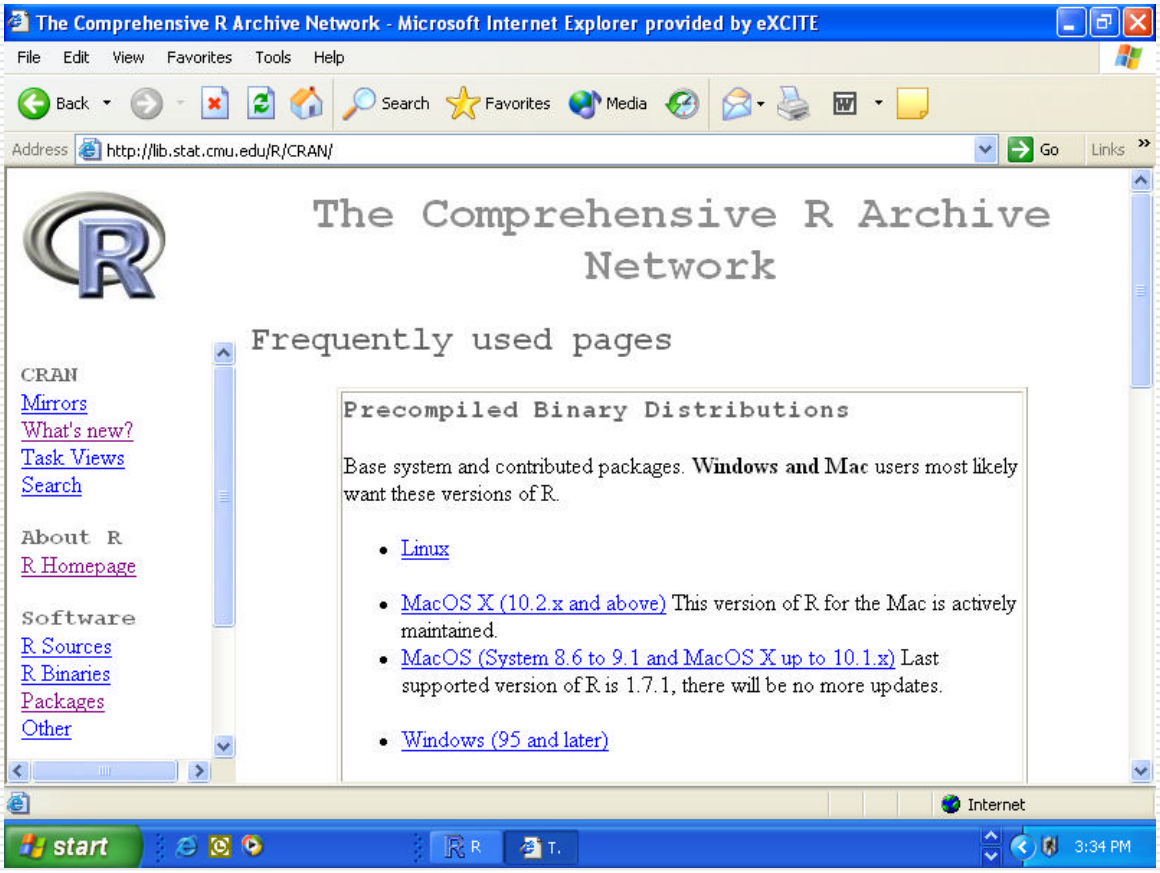
What about the outlier?

- Maximum 4.9 ppm
- Second Largest 1.8 ppm
- 95th percentile of US soil 2.2 ppm
- Largest is unlikely to be due to soil, should investigate

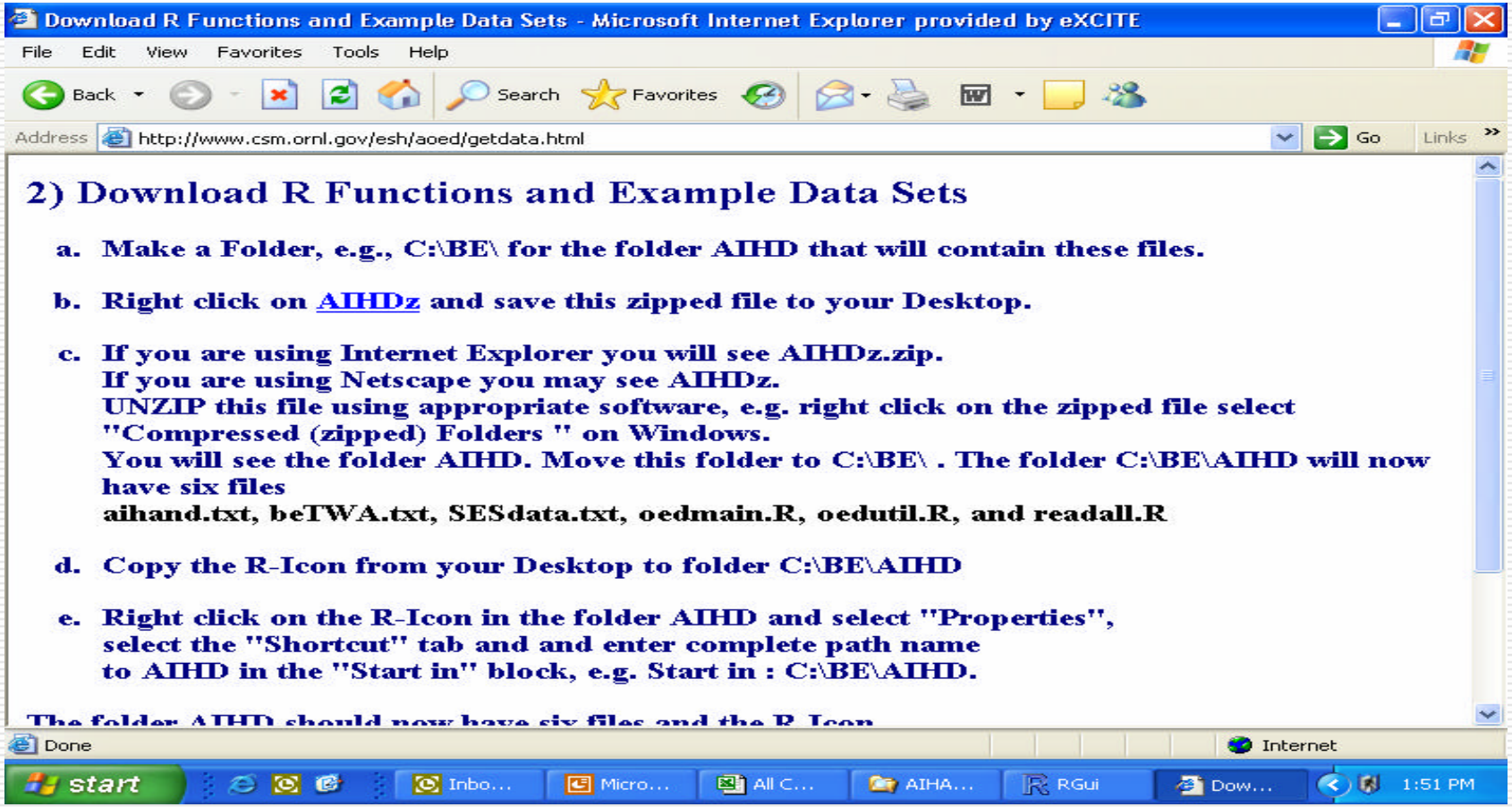
The R Foundation for Statistical Computing
http://www.r-project.org/



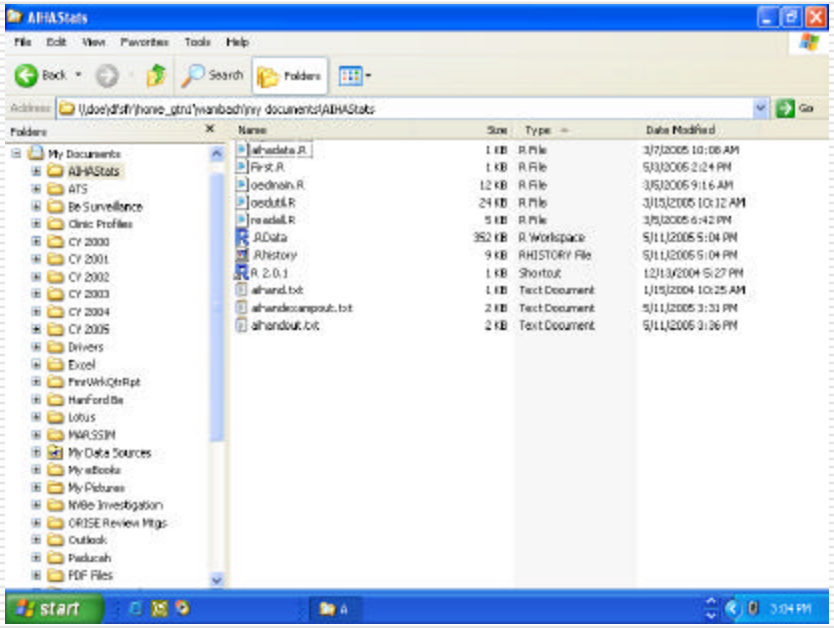
Download and Install R



http://www.csm.ornl.gov/esh/aoed/

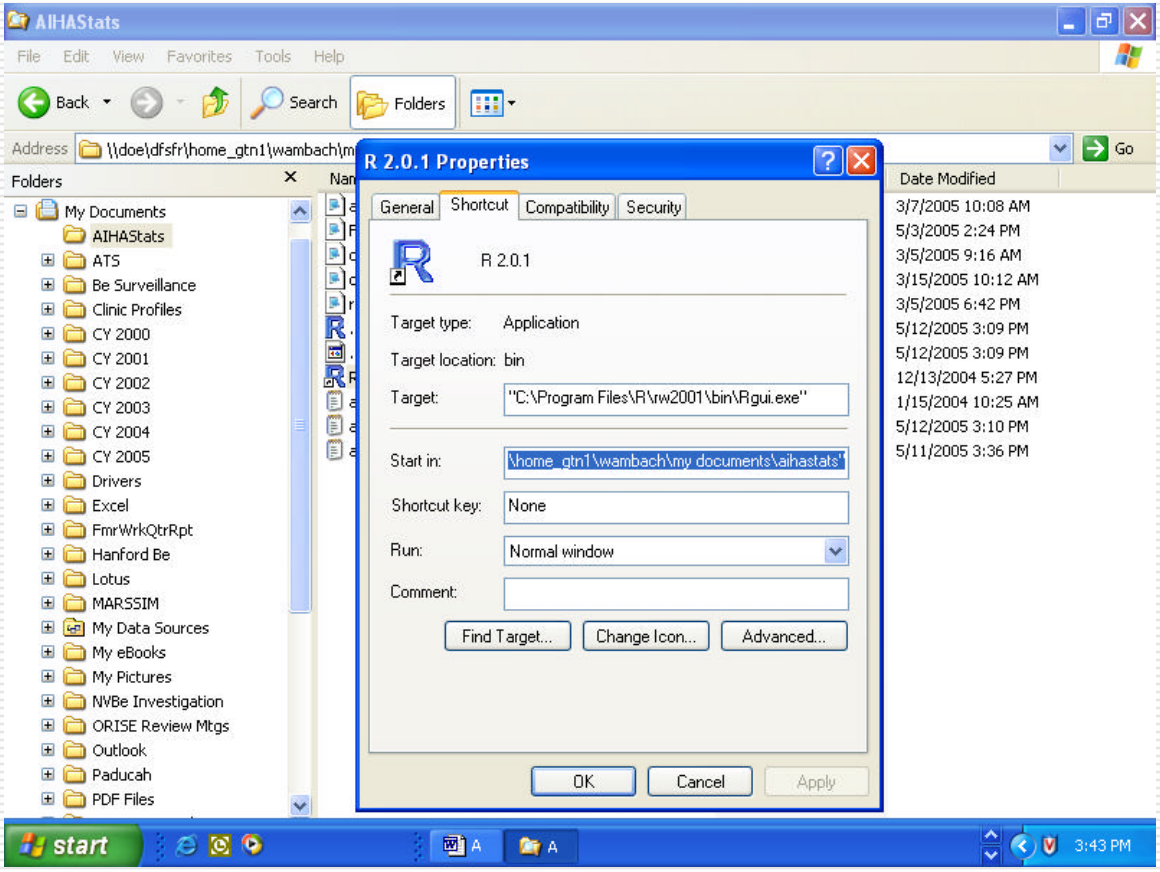


Set-up a Folder

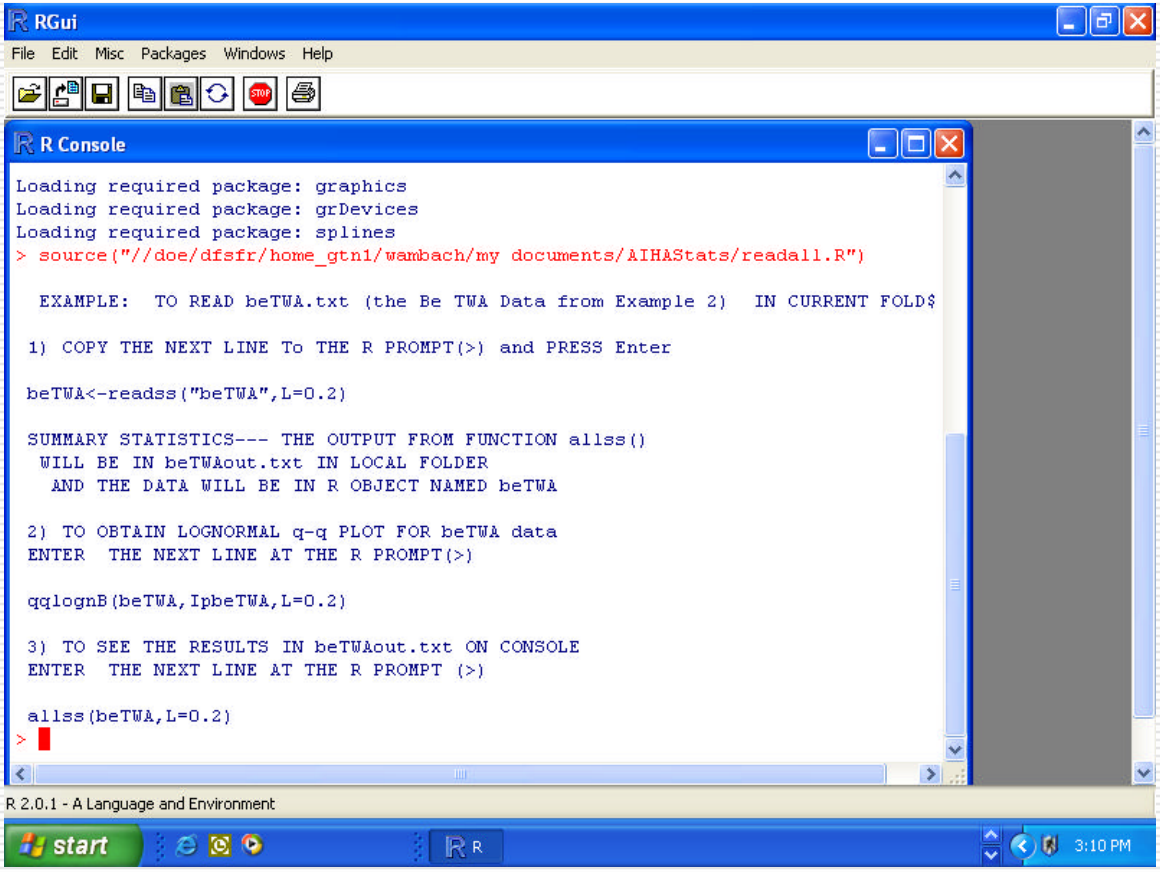


- Copy of R icon
- R Files saved from the web site
- Text files for analysis

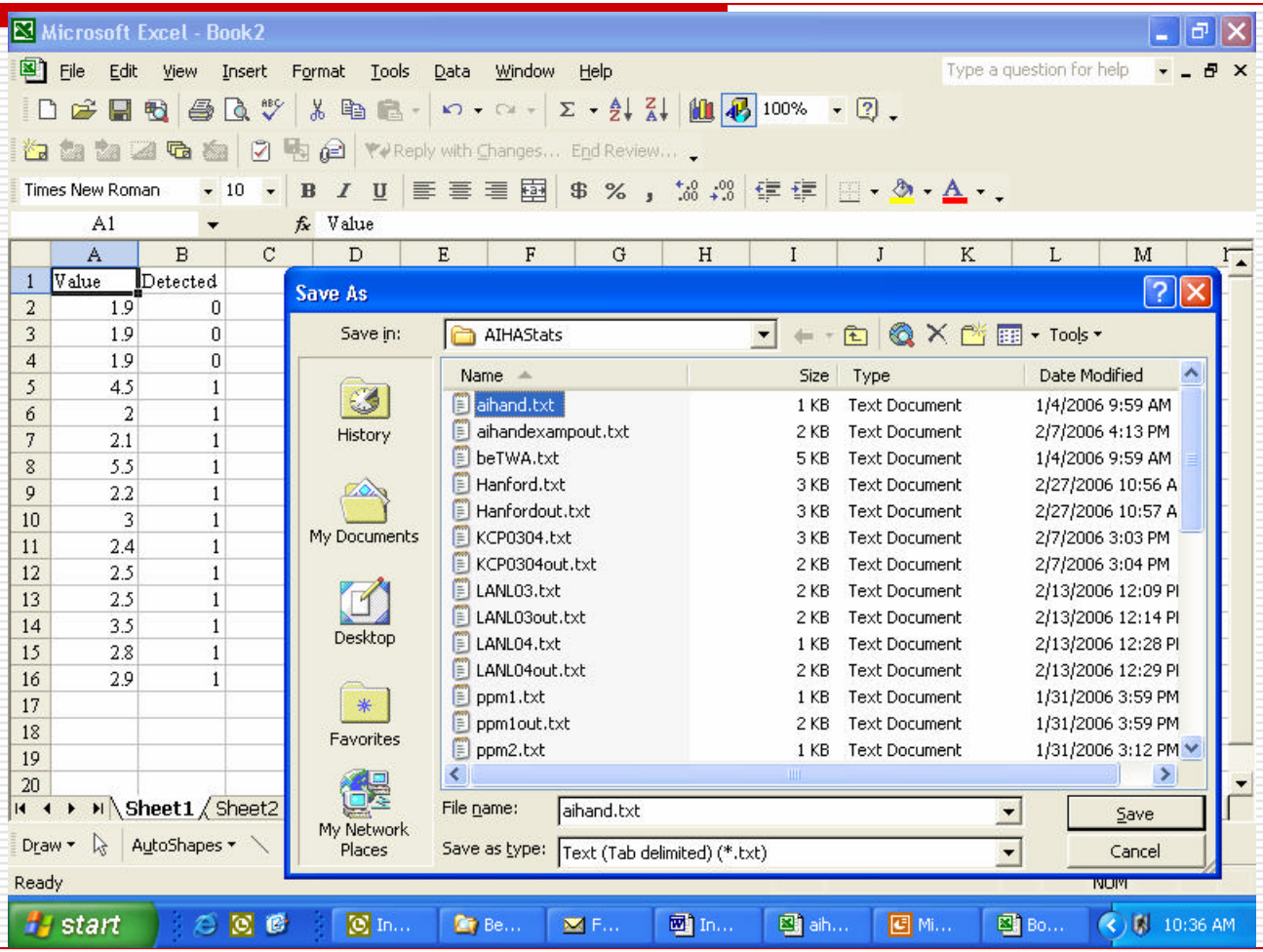
Set Icon Properties to Start In the Folder you Created



Open R, Select – File – Read Source Code – Readall.R



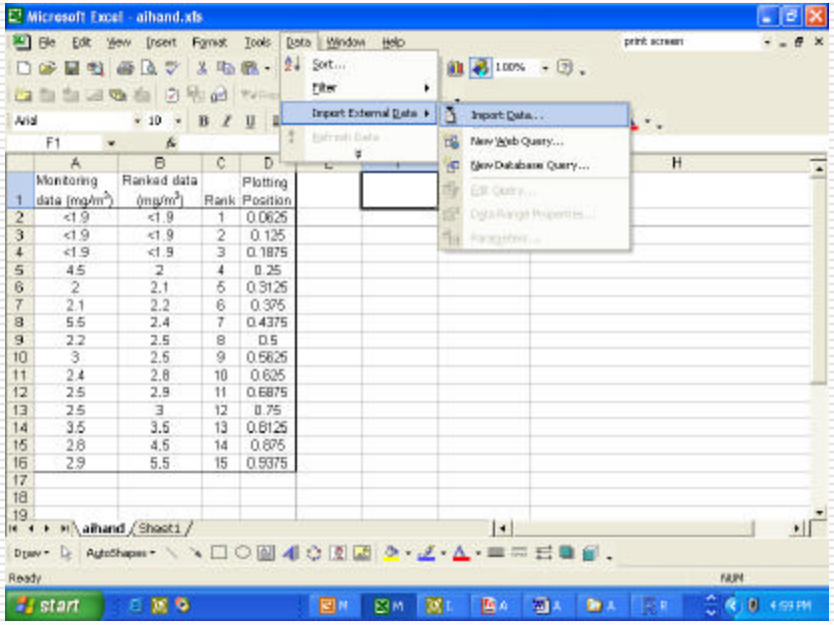
Save Data and Flag As Text (Tab delimited)(*.txt)

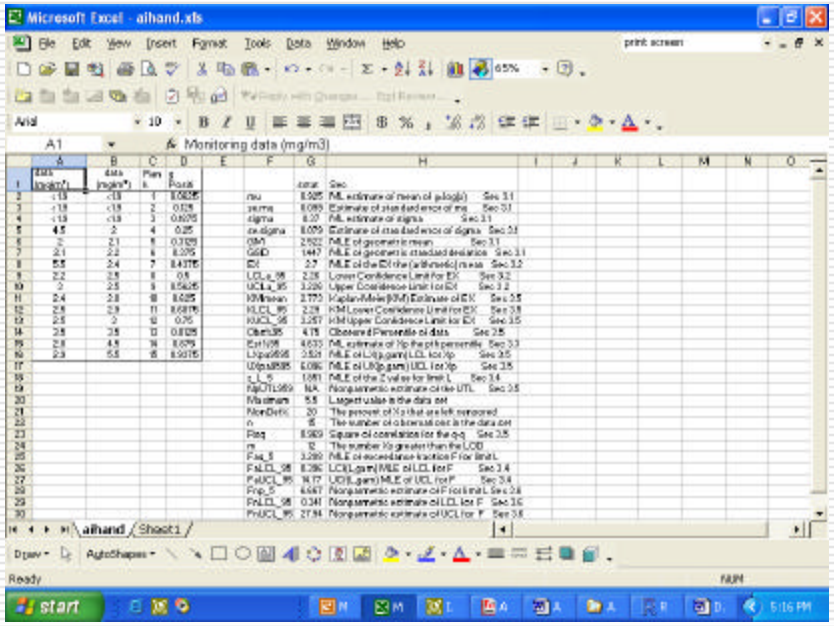


R Command
Filename<-readss(“Filename”,L=5)

- Filename is case sensitive
- L= the OEL being used to interpret the data
- Produces a text file with all metrics “Filenameout”

Import “Filenameout.txt”





R Command
allss(filename,L=5)

RGui - [R Console]

File Edit Misc Packages Windows Help

```
GM          2.522 MLE of geometric mean                  Sec 3.1
GSD         1.447 MLE of geometric standard deviation    Sec 3.1
EX          2.700 MLE of the EX the (arithmetic) mean    Sec 3.2
LCLa_95     2.260 Lower Confidence Limit for EX          Sec 3.2
UCLa_95     3.226 Upper Confidence Limit for EX          Sec 3.2
KMmean      2.773 Kaplan-Meier(KM) Estimate of EX        Sec 3.5
KLCL_95     2.290 KM Lower Confidence Limit for EX       Sec 3.5
KUCL_95     3.257 KM Upper Confidence Limit for EX       Sec 3.5
Obs%95      4.750 Obsevred Percentile of data            Sec 3.5
Est%95      4.633 ML estimate of Xp the pth percentile   Sec 3.3
LXpa9595    3.521 MLE of LX(p,gam) LCL for Xp            Sec 3.5
UXpa9595    6.096 MLE of UX(p,gam) UCL for Xp            Sec 3.5
z_L_5       1.851 MLE of the Z value for limit L         Sec 3.4
NpUTL9595      NA Nonparmetric estimate of the UTL       Sec 3.5
Maximum     5.500 Largest value in the data set
NonDet%    20.000 The percent of Xs that are left censored
n          15.000 The number of observations in the data set
Rsq         0.969 Square of correlation for the q-q      Sec 3.5
m          12.000 The number Xs greater than the LOD
Fax_5       3.208 MLE of exceedance fraction F for limit L
FaLCL_95    0.396 LCf(L,gam) MLE of LCL for F            Sec 3.4
FaUCL_95   14.767 UCf(L,gam) MLE of UCL for F            Sec 3.4
Fnp_5       6.667 Nonparmetric estimate of F for limit L Sec 3.6
FnLCL_95    0.341 Nonparmetric estimate of LCL for  F    Sec 3.6
FnUCL_95   27.940 Nonparmetric estimate of UCL for  F    Sec 3.6
>
```

R 2.0.1 - A Language and Environment

start R 3:24 PM

Contact Information

Paul Wambach, CIH

U.S. Department of Energy

Office of Occupational Health

Phone: 301-903-7373

Email:

Paul.Wambach@eh.doe.gov

Edward Frome, PhD

Oak Ridge National Laboratory

Computer Science and Mathematics Division

Phone: 865-574-3138

Email: FromeEL@ornl.gov

Web Site:

http://www.csm.ornl.gov/~frome/