Welcome to the CLU-IN Internet Seminar

ProUCL Webinar Part II

Sponsored by: USEPA ORD Site Characterization and Monitoring Technical Support Center (SCMTSC)

Delivered: March 16, 2011, 1:00 PM - 4:00 PM, EDT (17:00-20:00 GMT)

Instructors:

Anita Singh, PhD, Operations Researcher Senior Staff Scientist, Information System and Global Services - Civil, Lockheed Martin (asingh428@gmail.com)

Narian Armbya (narania@gmail.com)

Nancy Rios Jafolla, U.S. EPA Region 3 (rios-jafolla.nancy@epa.gov)

Moderator:

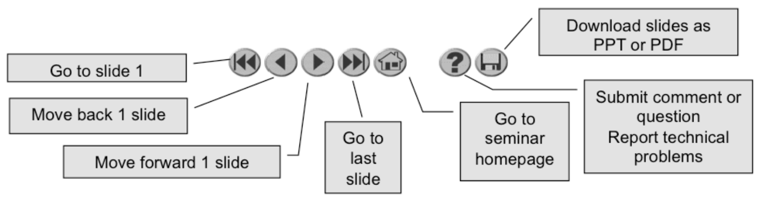
Felicia Barnett, U.S. EPA, ORD Site Characterization and Monitoring Technical Support Center (SCMTSC) (barnett.felicia@epa.gov)

1

Visit the Clean Up Information Network online at www.cluin.org

Housekeeping

- Please mute your phone lines, Do NOT put this call on hold
  - press *6 to mute #6 to unmute your lines at anytime (or applicable instructions)
- Q&A
- Turn off any pop-up blockers
- Move through slides using # links on left or buttons



- This event is being recorded
- Archives accessed for free **http://cluin.org/live/archive/**

2

Although I'm sure that some of you have these rules memorized from previous CLU-IN events, let's run through them quickly for our new participants.

Please mute your phone lines during the seminar to minimize disruption and background noise. If you do not have a mute button, press *6 to mute #6 to unmute your lines at anytime. Also, please do NOT put this call on hold as this may bring delightful, but unwanted background music over the lines and interupt the seminar.

You should note that throughout the seminar, we will ask for your feedback. You do not need to wait for Q&A breaks to ask questions or provide comments. To submit comments/questions and report technical problems, please use the ? Icon at the top of your screen. You can move forward/backward in the slides by using the single arrow buttons (left moves back 1 slide, right moves advances 1 slide). The double arrowed buttons will take you to 1st and last slides respectively. You may also advance to any slide using the numbered links that appear on the left side of your screen. The button with a house icon will take you back to main seminar page which displays our agenda, speaker information, links to the slides and additional resources. Lastly, the button with a computer disc can be used to download and save today's presentation materials.

With that, please move to slide 3.

ProUCL 4.1.00

Single and Two Sample Hypotheses Testing Approaches and Oneway ANOVA

http://www.epa.gov/osp/hstl/tsc/software.htm

3

Focus of ProUCL 4.1 Webinar II

- Focus of Webinar II is to make participants familiar with Statistical capabilities of ProUCL 4.1

- Emphasis will be placed on showing how to use ProUCL4.1 to:
  - Perform Single and Two Sample Hypotheses tests on data sets with nondetects (NDs) and without NDs
  - Perform Oneway Analysis of Variance (ANOVA)
  - Compute Background Threshold Values (BTVs)
  - Perform Trend Analysis using Linear Regression, Mann-Kendall trend test, and Theil-Sen trend test
  - Estimate mean, standard deviation, and Upper Limits based upon data sets with ND observations (e.g., KM method)
  - Interpret results generated by ProUCL

- Due to time limitation-statistical details will not be covered

LOCKHEED MARTIN

Null and Alternative Hypotheses

- H_0 : Null hypothesis statement (baseline condition)
- H_1 : Alternative hypothesis statement
  - Null hypothesis, H_0: Site mean $\leq C_s$
  - Alternative hypothesis, H_1: Site mean $> C_s$
- Hypotheses tests are performed on sampled data:
  - Therefore statistics used to test hypotheses suffer from uncertainties; and
  - Conclusions derived using those statistics suffer from decision errors

Decision Errors in Hypothesis Testing

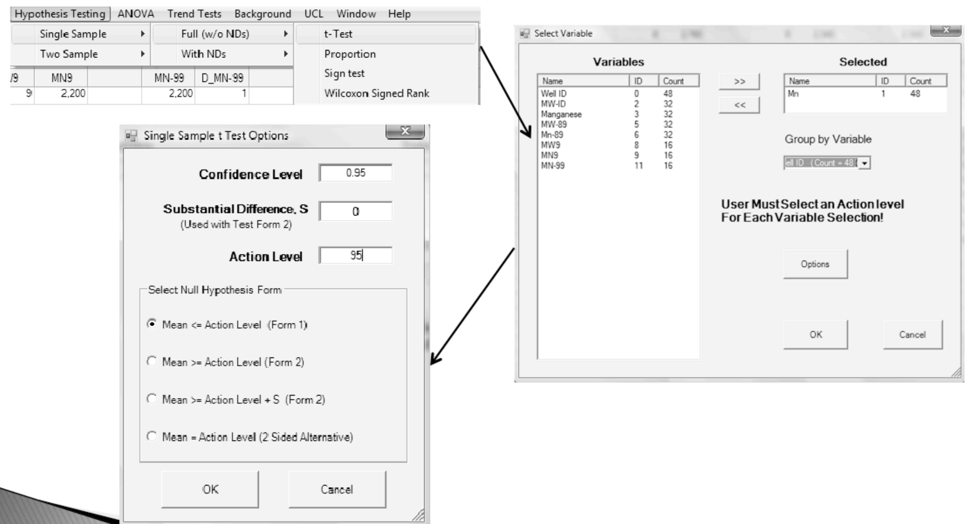
- Two types of decision errors can be made:
  - Type 1 Error = Probability (reject the null statement when it is true) = false positive error = α = false rejection rate
  - Type 2 Error = Probability (do not reject the null statement when it is false) = false negative error = β = false acceptance rate
- Width of gray region, α, and β are specified in DQOs
- Whenever possible, adequate amount of data should be collected based upon DQOs

LOCKHEED MARTIN

What is a P-value?

- P- value is associated with a test statistic such as a t-test
  - p-value is the smallest value of level of significance (Type I error) for which the null hypothesis is rejected
  - 1%, 5%, and 10% are common significance levels to which p-values are compared
    - A p-value < .05 rejects the null hypothesis at “ 5% level”
- ProUCL computes p-values for most of the hypothesis tests in ProUCL

Single Sample Hypothesis Tests in ProUCL



LOCKHEED MARTIN

Hypothesis Test to Verify Attainment of Clean Standard

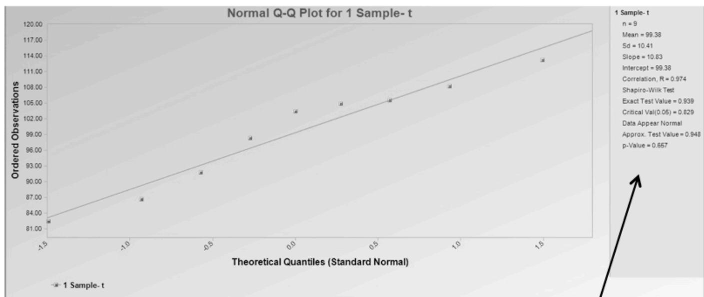
- Is site mean comparable to a cleanup threshold: C_s?
  - Null hypothesis, H_0: Site mean $\le C_s$
  - Alternative hypothesis, H_1: Site mean $> C_s$
- Use parametric or nonparametric test
  - Parametric t-test compares site mean with a threshold
  - Nonparametric tests :- Sign test and Wilcoxon Signed Rank (WSR) test compare site median with a threshold
- WSR test more powerful than Sign test

LOCKHEED MARTIN

T-Test: Compare Site Mean with Threshold

- 9 soil samples from a site area: 82.39 103.46 104.93 105.52 98.37 113.23 86.62 91.72 108.21 (EPA 2006)
  - Cleanup standard, $C_s = 95$

- Objective: Does site area meet cleanup standard?
  - H_0: Site mean ≤ 95 (meets standard), vs.
  - H_1: Site mean >95 (does not meet standard)



Data are normal, use parametric t-test

LOCKHEED MARTIN

T-Test: Compare Site Mean with Threshold

- ProUCL generated t-test results:

1 Sample- t

Single Sample t-Test

Raw Statistics

Number of Valid Observations	9
Number of Distinct Observations	9
Minimum	82.39
Maximum	113.2
Mean	99.38
Median	103.5
SD	10.41
SE of Mean	3.468

H0: Site Mean <= 95 (Form 1)

Test Value	1.264
Degrees of Freedom	8
Critical Value (0.05)	1.86
P-Value	0.121

Conclusion with Alpha = 0.05

Do Not Reject H0, Conclude Mean <= 95

P-Value > Alpha (0.05)

T-test statistic 1.26 < critical value 1.86
p-value = 0.12> 0.05

Conclusion based upon data:

Do not reject null hypothesis
Conclude : Site mean does not exceed 95

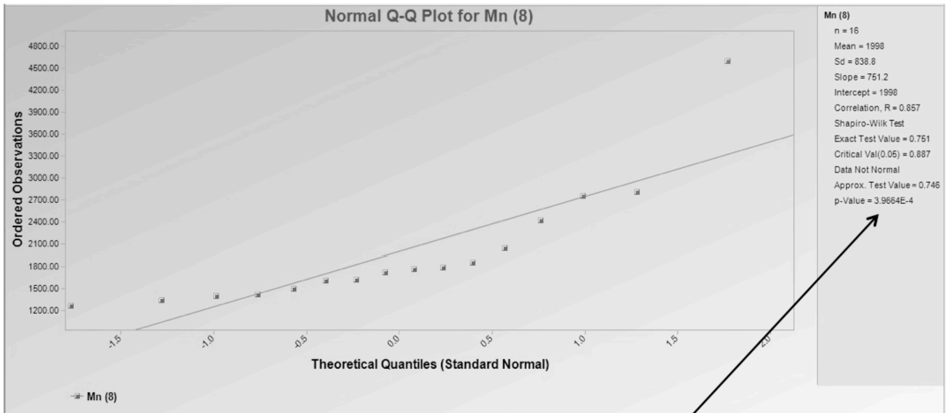
Sign Test: Compare Site Median with Threshold

- Sign test is used when data set is not normal and/or consists of NDs
- MW8 Mn data : 4600 2760 1270 1860 1790 1730 1420 1500 1610 1400 1350 1770 2050 2420 1630 2810
  - Cleanup threshold = 1500
  - H_0: MW8 median ≤ 1500 (threshold met)
  - H_1: MW8 median >1500
  - Data not normally distributed
  - Use nonparametric Sign test

12

LOCKHEED MARTIN

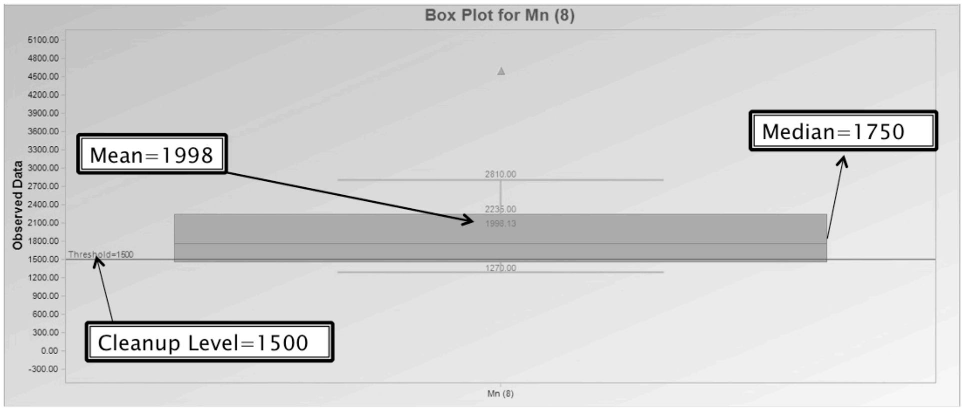
Sign Test: Compare Site Median with Threshold



P-value = 0.0003 (<0.05)
Data are not normal, use a nonparametric test

LOCKHEED MARTIN

Boxplot Comparing MW8 Median with Threshold



MW 8 mean = 1998, MW8 median = 1750, Threshold = 1500
Box plot suggests that site median exceeds threshold

LOCKHEED MARTIN

Sign Test: Compare Site Median with Threshold

- Sign test results:

Mn (8)

Single Sample Sign Test

Raw Statistics	
Number of Valid Observations	16
Number of Distinct Observations	16
Minimum	1270
Maximum	4600
Mean	1998
Median	1750
SD	838.8
SE of Mean	209.7
Number Above Action Level	11
Number Equal Action Level	1
Number Below Action Level	4

H0: Site Median <= 1500 (Form 1)

Test Value	11
P-Value	0.0592

Conclusion with Alpha = 0.05

Do Not Reject H0 at the specified level of significance (0.05). Conclude Median <= 1500

P-Value > Alpha (0.05)

p-value = 0.059 (>0.05)

Conclusion:

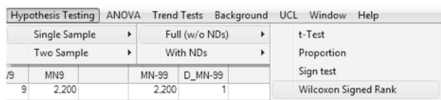
Do not reject H_0 for all levels of significance < 0.059

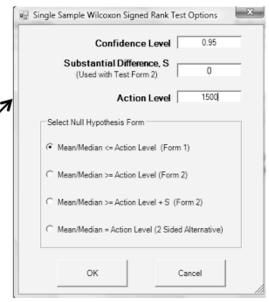
Based upon data and Sign test, conclude Mn in MW8 meets cleanup level of 1500 ppm at 0.05 level of significance

Is this correct conclusion?

WSR Test: Compare Site Median with Threshold

- Wilcoxon Signed Rank (WSR) test is more powerful than Sign test
  - WSR is used when data not normal and/or data consist of NDs
- Using MW8 Mn data:
  - H_0: MW8 median $\leq$ 1500 (MW8 meets threshold),
  - H_1: MW8 median $>$1500





WSR Test: Compare Site Median with Threshold

- WSR test results:

Mn (8)		
Single Sample Wilcoxon Signed Rank Test		
Raw Statistics		
Number of Valid Observations	16	
Number of Distinct Observations	16	
Minimum	1270	
Maximum	4600	
Mean	1998	
Median	1750	
SD	838.8	
SE of Mean	209.7	
Number Above Limit	11	
Number Equal Limit	1	
Number Below Limit	4	
T-plus	105.5	
T-minus	14.5	
H0: Site Median <= 1500 (Form 1)		
Test Value	105.5	
Critical Value (0.05)	101	
P-Value	0.0034	

H0: Site Median <= 1500 (Form 1)		
Test Value	105.5	
Critical Value (0.05)	101	
P-Value	0.0034	
Conclusion with Alpha = 0.05		
Reject H0, Conclude Mean/Median > 1500		
P-Value < Alpha (0.05)		

Reject H_0 at all levels > 0.0034

Conclusion: Based upon data and WSR test, conclude median Mn exceeds cleanup level of 1500ppm

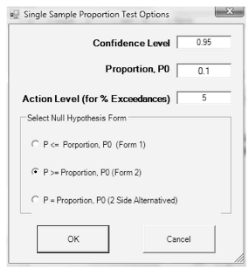
WSR test is more powerful than Sign test

Graphical display supports conclusion based upon WSR test

LOCKHEED MARTIN

Proportion Test to Compare Proportion of Exceedances with Allowable Proportion, P_0

- Used to determine if proportion of exceedances of an action level, A_0 by sampled data from a population (e.g., batch of drums, monitoring wells) meets pre-specified proportion, P_0 of exceedances
  - H_0: Proportion P of exceedances of A_0 by sampled data $\geq P_0$ vs.
  - H_1: Proportion P of exceedances of A_0 by sampled data $< P_0$
- If sample proportion p exceeds P_0 , population (e.g., lot of drums) is rejected requiring further investigation



Proportion Test to Compare Proportion of Exceedances with Allowable Proportion, P_0

- Sampled Data of size 85from EPA (2006): 4.19 5.3086 6.0524 3.3634 5.6631 5.0993 3.5597 5.8967 6.2773 4.9834 6.5021 7.3062 7.3321 5.505 7.4876 5.9948 7.1185 5.4988 6.1111 4.309 5.0479 3.9595 4.6125 5.6875 6.5491 7.6761 7.0345 6.8311 4.6146 6.6419 0.5981 5.898 5.7146 6.7668 5.5998 3.0195 5.2547 6.8017 4.0221 6.058 5.135 6.2445 6.0979 5.8625 3.6893 5.4765 5.5635 5.4628 6.0424 6.3631 5.88 5.89 1.46 4.05 1.09 2.59 1.69 3.16 2.08 2.61 3.42 2.54 4.91 4.1 6.74 7.27 7.42 7.5 6.56 4.64 5.98 3.14 3.23 5.8 6.17 6.01 5.8 3.6 5.765 5.55 5.48 3.693 5.9 5.5635 5.4
  - Action level, $A_0 = 5$
  - Allowable proportion, P_0 of exceedances = 0.1 (10%)
  - H_0: Proportion of exceedances of 5 by sampled data ≥ 0.1
  - H_1: Proportion of exceedances of 5 by sampled data < 0.1

Proportion Test to Compare Proportion of Exceedances with Allowable Proportion, P_0

- Proportion test results:

1Sample-Prop	
One-Sample Proportion Test	
Raw Statistics	
Number of Valid Observations	85
Number of Distinct Observations	83
Minimum	0.598
Maximum	7.676
Mean	5.183
Median	5.564
SD	1.588
SE of Mean	0.172
Number of Exceedances	56
Sample Proportion of Exceedances	0.659
H0: Site Proportion >= 0.1 (Form 2)	
Large Sample z-Test Value	16.99
Critical Value (0.05)	-1.645
P-Value	1
Conclusion with Alpha = 0.05	
Do Not Reject H0, Conclude Site Proportion >= 0.1	

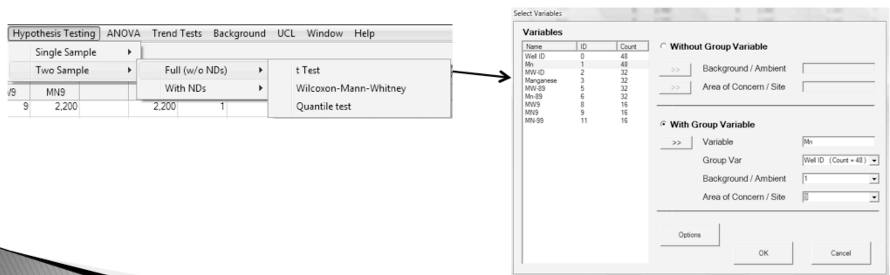
p-value=1
Null hypothesis, H_0 not rejected

Conclusion:
Proportion of exceedances
of action level, 5 exceeds 10%

Two Sample Hypothesis Tests

- Are site concentrations greater than background?
  - H_0: Site mean $\geq$ Background mean, vs.
  - H_1: Site mean $<$ Background mean

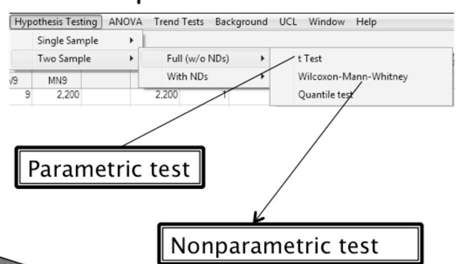
- Are subsurface soil concentrations comparable to surface soil concentrations?
  - H_0: Subsurface soil median $\leq$ Surface soil median, vs.
  - H_1: Subsurface soil median concentration $>$ Surface soil

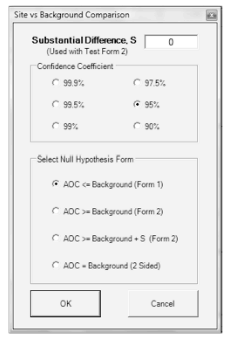


LOCKHEED MARTIN

Two Sample Hypothesis Tests

- Are downgradient MW concentrations of a COPC comparable to upgradient well concentrations?
  - H_0: MW mean ≤ Background well mean
  - H_1: MW mean > Background well mean
- Depending upon data distributions– parametric or nonparametric tests are used





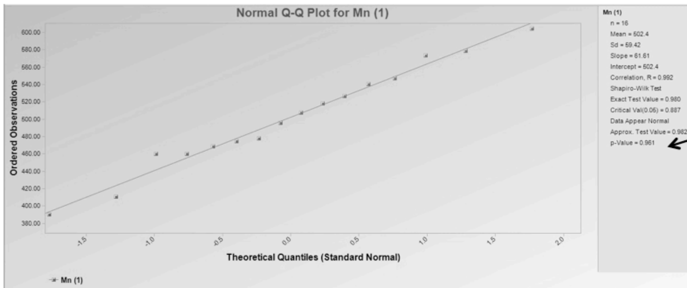
22

LOCKHEED MARTIN

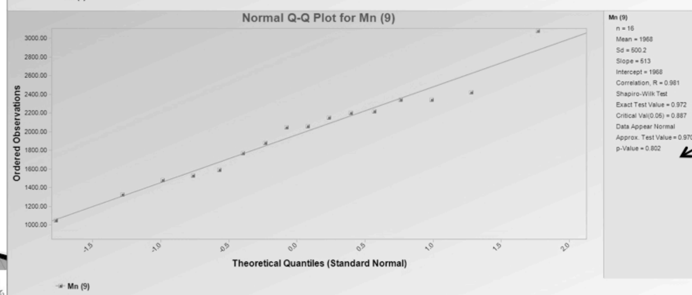
T-Test: Compare Manganese in Wells MW9 and Background Well MW1

- Mn in MW9:2200 2340 2340 2420 2150 2220 2050 2060 1770 1330 1590 1530 1480 1050 3080 1880
- Mn in Upgradient well, MW1:460 527 579 541 518 574 460 547 605 496 478 508 469 475 411 390
  - H_0: Mean Mn in MW9 $\leq$ Mean Mn in MW1
  - H_1: Mean of MW9 $>$ Mean of MW1
- T-test requires data sets to be normally distributed

T-Test: Compare Manganese in Wells: MW9 and Upgradient Well MW1



p-value =0.98
Mn in MW1
normal

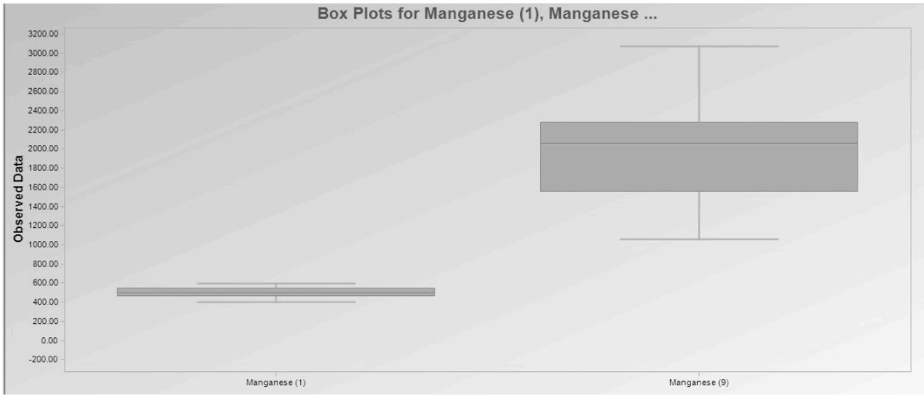


p-value=0.802
Mn in MW9
Normal

24

LOCKHEED MARTIN

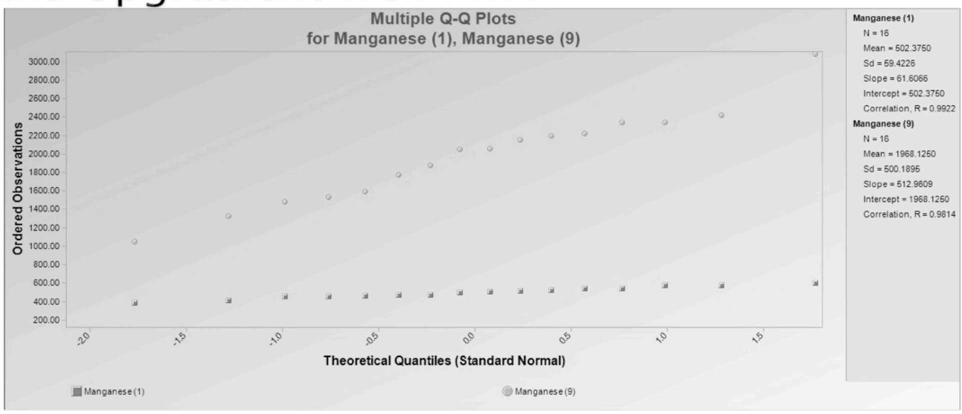
T-Test: Compare Manganese in Wells: MW9 and Upgradient Well MW1



Side-by-Side box plots comparison suggests that Mn in MW 9 > Mn in MW1

LOCKHEED MARTIN

T-Test: Compare Manganese in Wells: MW9 and Upgradient Well MW1



Q-Q plot comparison suggests that
Mn in MW 9 > Mn in MW1

LOCKHEED MARTIN

T-Test: Compare Manganese in Wells: MW9 and Upgradient Well MW1

Area of Concern Data: Manganese(9)

Background Data: Manganese(1)

Raw Statistics

	Site	Background
Number of Valid Observations	16	16
Number of Distinct Observations	15	15
Minimum	1050	390
Maximum	3080	605
Mean	1968	502.4
Median	2055	502
SD	500.2	59.42
SE of Mean	125	14.86

Site vs Background Two-Sample t-Test

H0: Mu of Site - Mu of Background <= 0

Method	DF	t-Test Value	Critical t (0.050)	P-Value
Pooled (Equal Variance)	30	11.640	1.697	0.000
Welch-Satterthwaite (Unequal Varia	15.4	11.640	1.753	0.000

Pooled SD 356.175

Conclusion with Alpha = 0.050

H0: Mu of Site - Mu of Background <= 0

Method	DF	t-Test Value	Critical t (0.050)	P-Value
Pooled (Equal Variance)	30	11.640	1.697	0.000
Welch-Satterthwaite (Unequal Varia	15.4	11.640	1.753	0.000

Pooled SD 356.175

Conclusion with Alpha = 0.050

* Student t (Pooled) Test: Reject H0, Conclude Site > Background
* Welch-Satterthwaite Test: Reject H0, Conclude Site > Background

Test of Equality of Variances

Variance of Site	250190
Variance of Background	3531

Numerator DF	Denominator DF	F-Test Value	P-Value
15	15	70.854	0.000

Conclusion with Alpha = 0.05

* Two variances are not equal

Conclusion

p-value = 0.0, Reject H_0
Conclude - Mean Mn in MW 9 is significantly higher than mean Mn in upgradient MW 1

Wilcoxon Mann Whitney (WMW) Test: Compare Onsite vs. Background Median Lead

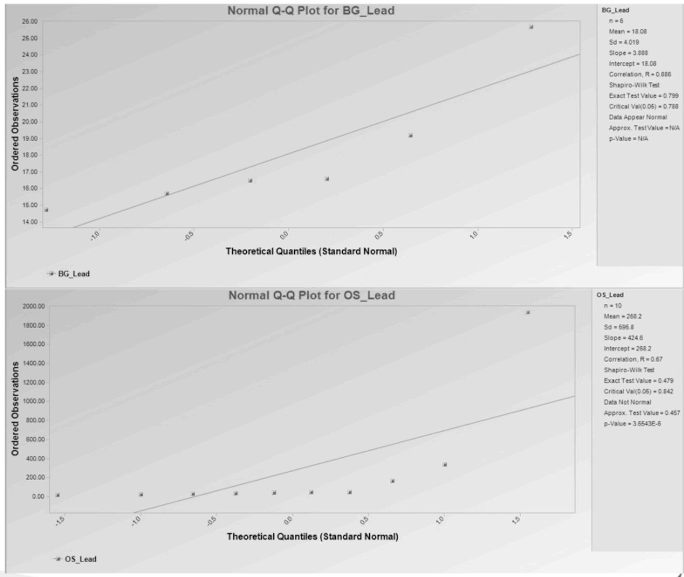
- Nonparametric WMW test can also be used on data sets with nondetects with a single detection limit

- Onsite and background data from a Superfund Site:
  - Onsite Lead: 27.1 38 23.8 38.6 19.7 47.4 165 338 1940 44.65
  - Background Lead: 25.7 15.7 16.6 16.5 14.75 19.2
  - H_0: Onsite median lead $\leq$ Background median lead, vs.
  - H_1: Onsite median lead $>$ background median lead concentration

WMW Test: Compare Onsite vs. Background Median Lead Concentrations

Background lead data not normal

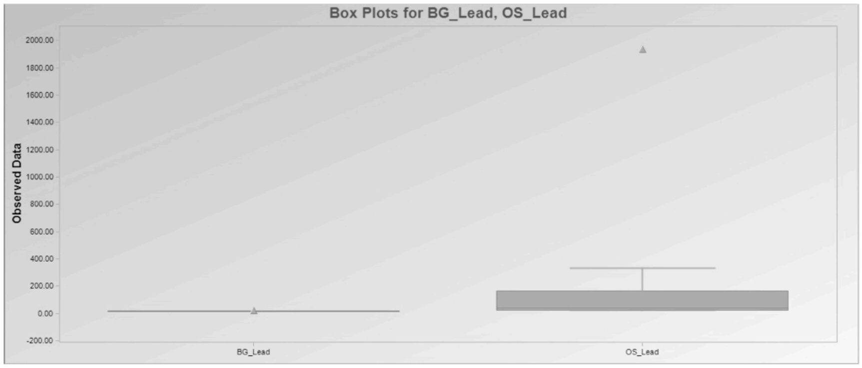
Onsite lead data not normal



LOCKHEED MARTIN

29

WMW Test: Compare Onsite vs. Background Median Lead Concentrations



Side-by-Side box plots comparison suggests that onsite lead > background lead

WMW Test: Compare Onsite vs. Background Median Lead Concentrations

Area of Concern Data: OS_Lead

Background Data: BG_Lead

Raw Statistics

	Site	Background
Number of Valid Observations	10	6
Number of Distinct Observations	10	6
Minimum	19.7	14.75
Maximum	1940	25.7
Mean	268.2	18.08
Median	41.63	16.55
SD	595.8	4.019
SE of Mean	188.4	1.641

Wilcoxon-Mann-Whitney (WMW) Test

H0: Mean/Median of Site or AOC <= Mean/Median of Background

Site Rank Sum W-Stat	113
WMW Test U-Stat	58
WMW Critical Value (0.050)	45
Approximate P-Value	0.00143

Conclusion with Alpha = 0.05

Reject H0, Conclude Site > Background

p-value = 0.0014
Reject null hypothesis & conclude:

Onsite median lead > Background median lead concentration

Conclusion

31

LOCKHEED MARTIN

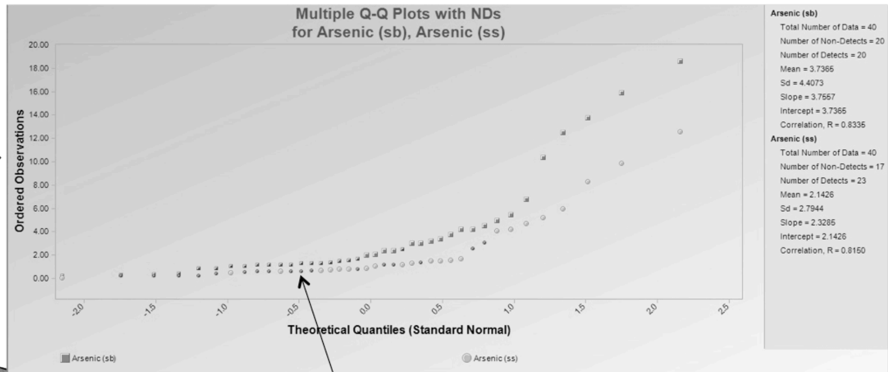
Gehan Test for Data Sets with NDs and Multiple Detection Limits (DLs)

- Nonparametric Gehan test can be used to compare concentrations of two populations when data sets consist of NDs with multiple DLs.
  - H_0: Surface soil median arsenic = subsurface soil median arsenic, vs.
  - H_1: Surface soil median arsenic ≠ subsurface soil median arsenic concentration

LOCKHEED MARTIN

Gehan Test on Arsenic Data with NDs and Multiple DLs

- Background arsenic data from subsurface (SB) and surface soils (SS) of a Federal Facility
- Need to estimate BTV
- Data set has multiple NDs exceeding detected values



LOCKHEED MARTIN

Gehan Test on Arsenic Data with NDs and Multiple DLs

Confidence Coefficient	95%
Substantial Difference	0.000
Selected Null Hypothesis	Site or AOC Mean/Median Equal to Background Mean/Median (Two Sided Alternative)
Alternative Hypothesis	Site or AOC Mean/Median Not Equal to Background Mean/Median

Area of Concern Data: Arsenic(ss)
Background Data: Arsenic(sb)

Raw Statistics

	Site	Background
Number of Valid Data	40	40
Number of Non-Detect Data	17	20
Number of Detect Data	23	20
Minimum Non-Detect	0.23	0.28
Maximum Non-Detect	3.1	2.5
Percent Non detects	42.50%	50.00%
Minimum Detected	0.094	2
Maximum Detected	12.6	18.6
Mean of Detected Data	3.048	6.335
Median of Detected Data	1.5	4.2
SD of Detected Data	3.369	5.038

Objective: Are surface and subsurface arsenic data comparable?

H_0: SS median=SUB median
H_1: SS median≠SUB median

H0: Mu of Site or AOC = Mu of background

Gehan z Test Value	-1.536
Lower Critical z (0.025)	-1.96
Upper Critical z (0.975)	1.96
P-Value	0.124

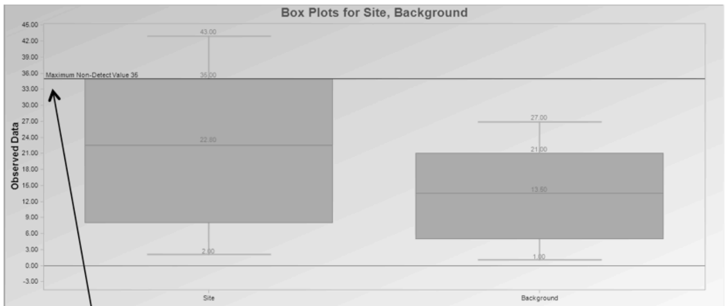
Conclusion with Alpha = 0.05
Do Not Reject H0, Conclude Site = Background
P-Value >= alpha (0.05)

P-value = 0.124
Arsenic in SS and SB not significantly different
Single estimate of BTV may be computed using combined surface and subsurface background data

LOCKHEED MARTIN

WMW and Gehan Tests on Data Sets with NDs and Multiple DLs

Data sets have multiple DLs with NDs > detects



Site data has highest ND > detected values
Box Plots suggest that Site arsenic > Background arsenic

LOCKHEED MARTIN

WMW Test on Data Sets with NDs and Multiple DLs

Area of Concern Data: Site
Background Data: Background

Raw Statistics

	Site	Background
Number of Valid Data	10	10
Number of Non-Detect Data	2	4
Number of Detect Data	8	6
Minimum Non-Detect	4	4
Maximum Non-Detect	35	25
Percent Non detects	20.00%	40.00%
Minimum Detected	2	1
Maximum Detected	43	27
Mean of Detected Data	23.63	12.17
Median of Detected Data	22.5	11
SD of Detected Data	14.74	9.642

Wilcoxon-Mann-Whitney Site vs Background Test
All observations <= 35 (Max DL) are ranked the same
Wilcoxon-Mann-Whitney (WMW) Test

H0: Mean/Median of Site or AOC <= Mean/Median of Background

Site Rank Sum W-Stat	115
WMW Test U-Stat	60

H0: Mean/Median of Site or AOC <= Mean/Median of Background

Site Rank Sum W-Stat	115
WMW Test U-Stat	60
WMW Critical Value (0.050)	72
Approximate P-Value	0.236

Conclusion with Alpha = 0.05
Do Not Reject H0, Conclude Site <= Background

H_0: Site median≤ Background median
H_1: Site median> Background median

P-value = 0.236, Do not Reject H_0

WMW test incorrectly concludes:

Site median is comparable to background median

LOCKHEED MARTIN

Gehan Test on Data Sets with NDs and Multiple DLs

Raw Statistics

	Site	Background
Number of Valid Data	10	10
Number of Non-Detect Data	2	4
Number of Detect Data	8	6
Minimum Non-Detect	4	4
Maximum Non-Detect	35	25
Percent Non detects	20.00%	40.00%
Minimum Detected	2	1
Maximum Detected	43	27
Mean of Detected Data	23.63	12.17
Median of Detected Data	22.5	11
SD of Detected Data	14.74	9.642

Site vs Background Gehan Test

H0: Mean/Median of Site or AOC <= Mean/Median of background

Gehan z Test Value	1.769
Critical z (0.95)	1.645
P-Value	0.0384

Conclusion with Alpha = 0.05

Reject H0, Conclude Site > Background

P-Value < alpha (0.05)

H_0: Site median≤ Background median
H_1: Site median> Background median

P-value = 0.038, Reject H_0

Gehan test correctly concludes:
Site median > Background median

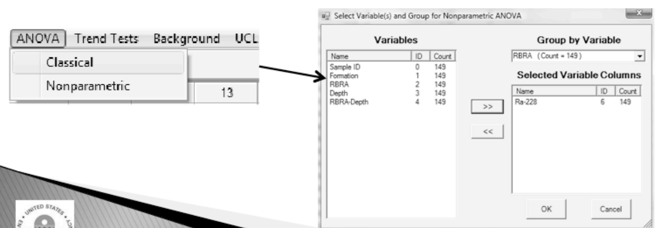
It is suggested to confirm test results with graphical displays

Oneway Analysis of Variance (ANOVA)

- Oneway ANOVA is used to compare mean/median concentrations of more than two groups such as:
  - Arsenic concentrations in several AOCs
  - Inter-well comparisons

  - Null Hypothesis: Mean concentrations are similar
  - Alternative hypothesis: Mean concentrations are different

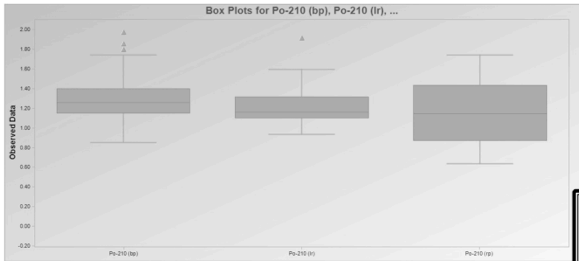
- Classical and nonparametric Kruskal-Wallis(K-W) ANOVA:

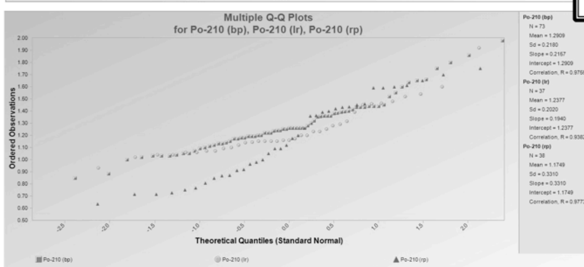


Oneway ANOVA on Po-210 data from 3 Areas

- Consider Polonium -210 (Po-210) data from 3 reference areas (LR, RP, and BP).
- Objective: Compute site-specific estimate of background level concentration for Po-210.
- First determine if Po-210 data from 3 areas come from the same population
- H_0: Po-210 means of 3 areas are comparable, vs.
- H_1: Po-210 means of 3 areas are not comparable

Oneway ANOVA on Po -210 data from 3 Areas





Graphical displays suggest that 3 areas exhibit comparable Po-210 concentrations

LOCKHEED MARTIN

Oneway ANOVA on Po -210 data from 3 Areas

	Classical One-Way ANOVA			
Date/Time of Computation	2/16/2011 5:45:18 PM			
From File	ANOVA			
Full Precision	OFF			
Po-210				
Group	Obs	Mean	SD	Variance
lr	37	1.238	0.202	0.0408
rp	38	1.175	0.331	0.11
bp	73	1.291	0.218	0.0475
Grand Statistics (All data)	148	1.248	0.251	0.0632

Classical One-Way Analysis of Variance Table					
Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	0.342	2	0.171	2.768	0.0661
Within Groups	8.944	145	0.0617		
Total	9.286	147			
Pooled Standard Deviation		0.248			
R-Sq		0.0368			

P-value = 0.066
Null hypothesis not rejected

Conclude:
Po-210 data from 3 areas come from same background population

A single estimate of background threshold can be computed based upon merged Po-210 data set

Nonparametric K-W Oneway ANOVA on Mn Data

- Consider Mn data from 3 monitoring wells (MWs)
- Objective: Perform Inter-well comparison
  - H_0: Median Mn concentrations of 3 MWs are comparable, vs.
  - H_1: Median Mn concentrations of 3 areas are not comparable
- Kruskal-Wallis Test results are shown next

Nonparametric K-W Oneway ANOVA on Mn data

	Nonparametric One Way ANOVA (Kruskal-Wallis Test)				
User Selected Options					
Date/Time of Computation	2/16/2011 6:20:36 PM				
From File	ANOVA_NP				
Full Precision	OFF				
Mn					
Group	Obs	Median	Ave Rank	Z	
1	16	502	8.5	-5.599	
8	16	1750	31.41	2.417	
9	16	2055	33.59	3.182	
Overall	48	1515	24.5		
K-W (H-Stat)	DOF	P-Value	(Approx. Chisquare)		
31.54	2	1.4148E-7			
31.55	2	1.4087E-7	(Adjusted for Ties)		

p-value = 0.0
Null hypothesis rejected at all levels

Conclude:

Mn in 3 MWs are significantly different

43

LOCKHEED MARTIN

ProUCL 4.1.00

Upper Percentile, Upper Prediction Limit (UPL), and Upper Tolerance Limit (UTL) to Estimate Background Threshold Values (BTVs)

http://www.epa.gov/osp/hstl/tsc/software.htm

44

Upper Percentiles, UPLs, and UTLs

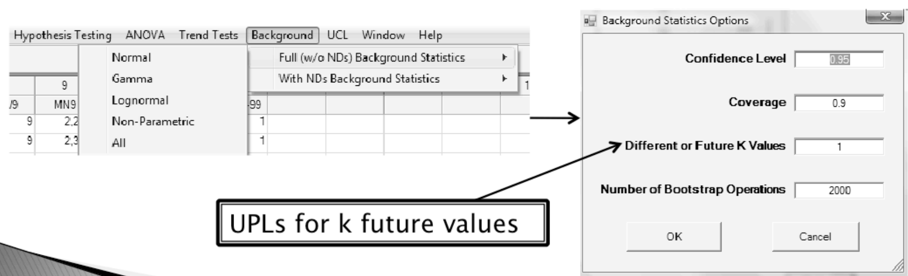
- Based upon an "established background data set", upper percentiles, UPL95, and UTL95-95 are computed to:
  - Estimate background level concentrations, background threshold values (BTVs), not-to-exceed values
- Onsite observations are compared with BTVs to:
  - Identify contaminants of potential concern
  - Determine potentially polluted site locations
  - Perform Intra-well comparisons to identify non-complying wells in groundwater (GW) studies

LOCKHEED MARTIN

UPLs and UTLs – "Full" Data without NDs

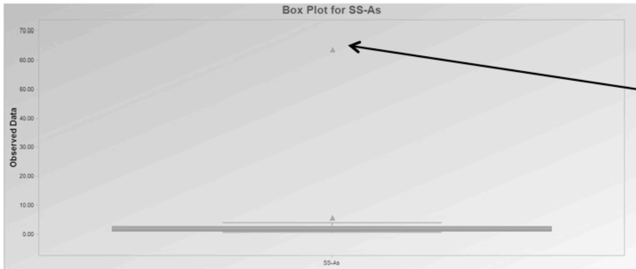
- Surface soil arsenic background data of size 39 from a Federal Facility: 1.9 1.2 1.5 3.4 3.6 2 5.8 1.4 1.5 1.2 2.9 0.46 0.65 0.84 0.75 0.53 2.5 2.5 4.2 2.9 63.9 1.8 1.1 1.5 1.4 1.1 2.7 2.9 1.2 2.4 1.1 1.4 4.1 2.1 0.68 1.6 1.7 0.83 1.2

- Estimate BTVs:
  - Are there any outliers?
  - Data normal, lognormal or gamma distribution?

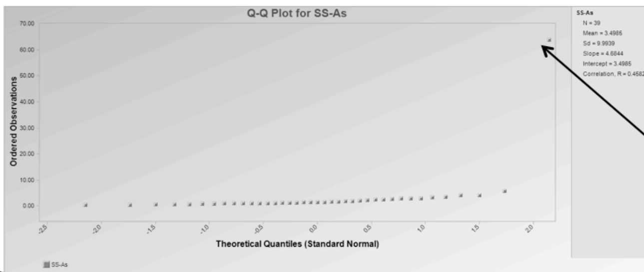


LOCKHEED MARTIN

Graphical Displays



Boxplot – there is at least one potential outlier



Q-Q plot – there is at least one potential outlier

47

Rosner Outlier Test (n>25)

Rosner's Outlier Test for SS-As

Mean	**3.498**
Standard Deviation	**9.994**
Number of data	**39**
Number of suspected outliers	**3**

#	Mean	sd	Potential outlier	Obs. Number	Test value	Critical value (5%)	Critical value (1%)
1	3.498	9.865	63.9	21	6.123	3.03	3.37
2	1.909	1.174	5.8	7	3.315	3.01	3.36
3	1.804	0.992	4.2	19	2.415	3	3.34

For 5% significance level, there are 2 Potential Outliers

Therefore, Potential Statistical Outliers are

63.9, 5.8

For 1% Significance Level, there is 1 Potential Outlier

Therefore, Observation 63.9 is a Potential Statistical Outlier

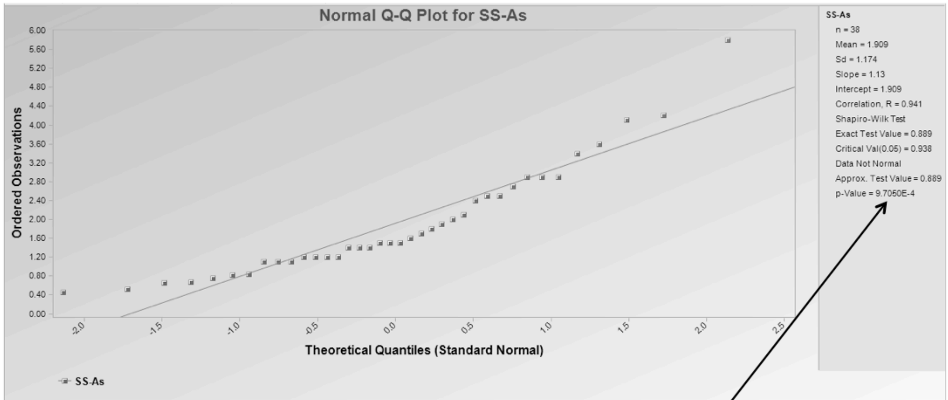
At 0.05 level of significance: 63.9 is high outlier, 5.8 is mild outlier

At 0.01 level of significance: 63.9 is the only outlier

Project Team should determine if 5.8 represents an outlier

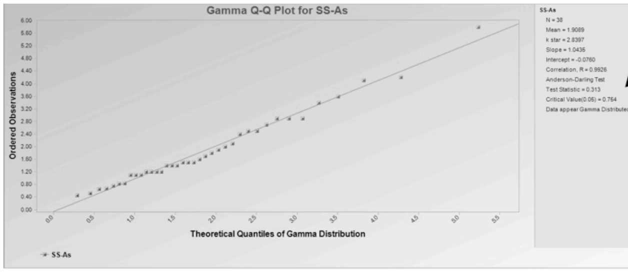
Here only 63.9 is considered outlier, BTVs are estimated without 63.9

GOF Test without Outlier 63.9



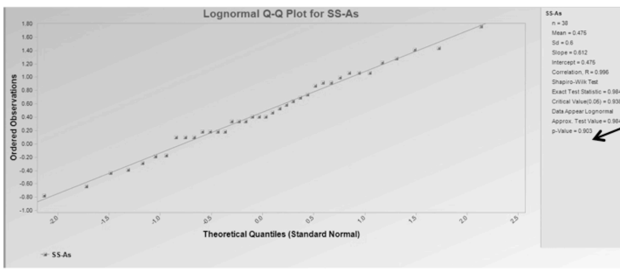
p-value for S-W test = 0.001<0.05
Conclude: Data not normally distributed

GOF Test without Outlier 63.9



Test value= 0.313 <
Critical Value =0.754

Data follow gamma
distribution at 0.05
level of significance



p-value = 0.9 for S-W Test
Data follow lognormal
distribution

LOCKHEED MARTIN

BTV Estimates Less Outlier 63.9 (All Option)

Background Statistics			
Normal Distribution Test		**Lognormal Distribution Test**	
Shapiro Wilk Test Statistic	0.889	Shapiro Wilk Test Statistic	0.984
Shapiro Wilk Critical Value	0.938	Shapiro Wilk Critical Value	0.938
Data not Normal at 5% Significance Level		**Data appear Lognormal at 5% Significance Level**	
Assuming Normal Distribution		**Assuming Lognormal Distribution**	
95% UTL with 95% Coverage	4.412	95% UTL with 95% Coverage	5.778
95% UPL (t)	3.915	95% UPL (t)	4.483
90% Percentile (z)	3.413	90% Percentile (z)	3.468
95% Percentile (z)	3.84	95% Percentile (z)	4.313
99% Percentile (z)	4.64	99% Percentile (z)	6.493
Gamma Distribution Test		**Data Distribution Test**	
k star	2.84	**Data appear Gamma Distributed at 5% Significance Level**	
Theta Star	0.672		
MLE of Mean	1.909		
MLE of Standard Deviation	1.133		
nu star	215.8		
A-D Test Statistic	0.313	**Nonparametric Statistics**	
5% A-D Critical Value	0.754	90% Percentile	3.46
K-S Test Statistic	0.11	95% Percentile	4.115
5% K-S Critical Value	0.144	99% Percentile	5.208
Data appear Gamma Distributed at 5% Significance Level			

Lognormal percentiles

BTVs Estimates without Outlier 63.9

Gamma Distribution Test		**Data Distribution Test**	
k star	2.84	**Data appear Gamma Distributed at 5% Significance Level**	
Theta Star	0.672		
MLE of Mean	1.909		
MLE of Standard Deviation	1.133		
nu star	215.8		
A-D Test Statistic	0.313	**Nonparametric Statistics**	
5% A-D Critical Value	0.754	90% Percentile	3.46
K-S Test Statistic	0.11	95% Percentile	4.115
5% K-S Critical Value	0.144	99% Percentile	5.208
Data appear Gamma Distributed at 5% Significance Level			
Assuming Gamma Distribution		95% UTL with 95% Coverage	5.8
90% Percentile	3.428	95% Percentile Bootstrap UTL with 95% Coverage	5.8
95% Percentile	4.07	95% BCA Bootstrap UTL with 95% Coverage	5.8
99% Percentile	5.467	95% UPL	4.28
		95% Chebyshev UPL	7.092
95% WH Approx. Gamma UPL	4.122	Upper Threshold Limit Based upon IQR	4.563
95% HW Approx. Gamma UPL	4.186		
95% WH Approx. Gamma UTL with 95% Coverage	4.952		
95% HW Approx. Gamma UTL with 95% Coverage	5.098		

Gamma Percentiles

Estimate of BTV

	Gamma	Lognormal
UPL95	4.12	4.48
UTL95-95	4.95	5.78

Use of UPLs and UTLs based upon gamma distribution is suggested

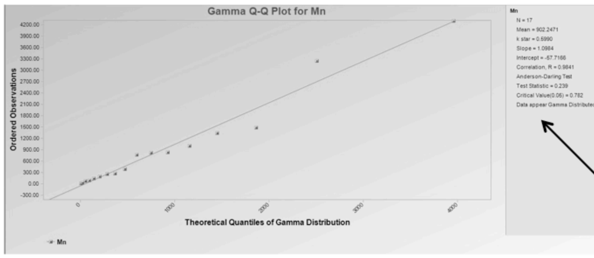
LOCKHEED MARTIN

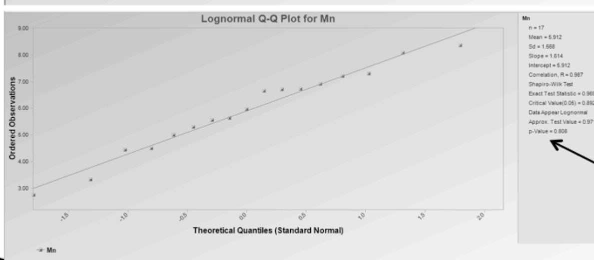
Influence of Outliers on UPLs and UTLs

- Mn data from a Navy Site:15.8 28.2 90.6 1490 85.6 281 4300 199 838 777 824 1010 1350 390 150 3250 259
  - As determined in UCL95 section, 4300 and 3250 represent potential outliers
    - Project team should make a decision about their disposition
  - Data (with outliers)follow gamma as well as lognormal distribution – see next slide
  - Just like UCL95, lognormal distribution based (with or without outliers) UPLs and UTLs are unacceptably large

GOF Tests on Mn Data with outliers

- Mn data from a Navy Site





Data gamma distributed

A-D Test statistic=0.239

Critical value (.05) =0.782 > 0.239

Data are lognormal p-value = 0.808

LOCKHEED MARTIN

Influence of Outliers on UPLs and UTLs

Raw Statistics		Log-Transformed Statistics	
Minimum	15.8	Minimum	2.76
Maximum	4300	Maximum	8.366
Second Largest	3250	Second Largest	8.086
First Quartile	150	First Quartile	5.011
Median	390	Median	5.966
Third Quartile	1010	Third Quartile	6.918
Mean	902.2	Mean	5.912
SD	1189	SD	1.568
Coefficient of Variation	1.318		
Skewness	2.046		

Background Statistics

Normal Distribution Test		Lognormal Distribution Test	
Shapiro Wilk Test Statistic	0.725	Shapiro Wilk Test Statistic	0.969
Shapiro Wilk Critical Value	0.892	Shapiro Wilk Critical Value	0.892
Data not Normal at 5% Significance Level		**Data appear Lognormal at 5% Significance Level**	
Assuming Normal Distribution		**Assuming Lognormal Distribution**	
95% UTL with 95% Coverage	3859	95% UTL with 95% Coverage	18203
95% UPL (t)	3039	95% UPL (t)	6176
90% Percentile (z)	2427	90% Percentile (z)	2755
95% Percentile (z)	2859	95% Percentile (z)	4869
99% Percentile (z)	3669	99% Percentile (z)	14173

Inflated by outliers

Statistics computed using 2 outliers

LOCKHEED MARTIN

Influence of Outliers on UPLs and UTLs

Gamma Distribution Test		**Data Distribution Test**	
k star	0.599	**Data appear Gamma Distributed at 5% Significance Level**	
Theta Star	1506		
MLE of Mean	902.2		
MLE of Standard Deviation	1166		
nu star	20.37		
A-D Test Statistic	0.239	**Nonparametric Statistics**	
5% A-D Critical Value	0.782	90% Percentile	2194
K-S Test Statistic	0.117	95% Percentile	3460
5% K-S Critical Value	0.218	99% Percentile	4132
Data appear Gamma Distributed at 5% Significance Level			
Assuming Gamma Distribution		95% UTL with 95% Coverage	4300
90% Percentile	2347	95% Percentile Bootstrap UTL with 95% Coverage	4300
95% Percentile	3249	95% BCA Bootstrap UTL with 95% Coverage	4300
99% Percentile	5428	95% UPL	4300
		95% Chebyshev UPL	6237
95% WH Approx. Gamma UPL	3423	Upper Threshold Limit Based upon IQR	2300
95% HW Approx. Gamma UPL	3688		
95% WH Approx. Gamma UTL with 95% Coverage	5595		
95% HW Approx. Gamma UTL with 95% Coverage	6508		

Statistics computed using 2 outliers

LOCKHEED MARTIN

UPLs and UTLs without 2 Outliers

Raw Statistics		Log-Transformed Statistics	
Minimum	15.8	Minimum	2.76
Maximum	1490	Maximum	7.307
Second Largest	1350	Second Largest	7.208
First Quartile	120.3	First Quartile	4.759
Median	281	Median	5.638
Third Quartile	831	Third Quartile	6.723
Mean	519.2	Mean	5.604
SD	491.4	SD	1.392
Coefficient of Variation	0.946		
Skewness	0.807		
Background Statistics			
Normal Distribution Test		Lognormal Distribution Test	
Shapiro Wilk Test Statistic	0.87	Shapiro Wilk Test Statistic	0.929
Shapiro Wilk Critical Value	0.881	Shapiro Wilk Critical Value	0.881
Data not Normal at 5% Significance Level		Data appear Lognormal at 5% Significance Level	
Assuming Normal Distribution		Assuming Lognormal Distribution	
95% UTL with 95% Coverage	1780	95% UTL with 95% Coverage	9666
95% UPL (t)	1413	95% UPL (t)	3417
90% Percentile (z)	1149	90% Percentile (z)	1616
95% Percentile (z)	1328	95% Percentile (z)	2681
99% Percentile (z)	1662	99% Percentile (z)	6924

Elevated lognormal estimates

Statistics computed without 2 outliers 4300 and 3250

UPLs and UTLs without 2 Outliers

Gamma Distribution Test		**Data Distribution Test**	
k star	0.765	**Data appear Gamma Distributed at 5% Significance Level**	
Theta Star	678.5		
MLE of Mean	519.2		
MLE of Standard Deviation	593.5		
nu star	22.96		
A-D Test Statistic	0.298	**Nonparametric Statistics**	
5% A-D Critical Value	0.768	90% Percentile	1214
K-S Test Statistic	0.175	95% Percentile	1392
5% K-S Critical Value	0.229	99% Percentile	1470
Data appear Gamma Distributed at 5% Significance Level			
Assuming Gamma Distribution		95% UTL with 95% Coverage	1490
90% Percentile	1276	95% Percentile Bootstrap UTL with 95% Coverage	1490
95% Percentile	1711	95% BCA Bootstrap UTL with 95% Coverage	1490
99% Percentile	2743	95% UPL	1490
		95% Chebyshev UPL	2731
95% WH Approx. Gamma UPL	1863	Upper Threshold Limit Based upon IQR	1897
95% HW Approx. Gamma UPL	2039		
95% WH Approx. Gamma UTL with 95% Coverage	3015		
95% HW Approx. Gamma UTL with 95% Coverage	3556		

Statistics computed without 2 outliers
4300 and 3250

LOCKHEED MARTIN

95% UPLs with and without 2 outliers

- Outliers (4300, 3250) distorted all statistics
- Lognormal distribution yields unrealistically high values
- Data are gamma distributed
  - Use of UPLs or UTLs based upon gamma distribution is suggested

Method	With outliers n=17	Without outliers n=15
Normal UPL95	3039	1413
Gamma UPL95 (WH)	3423	1863
Lognormal UPL95	6176	3417
Nonparametric UPL95	4300	1490
Nonparametric 95% Percentile	3460	1392
Normal 95% Percentile	2859	1328
Maximum	4300	1490

Elevated values

LOCKHEED MARTIN

95-95 UTLs with and without 2 Outliers

- Outliers (4300 and 3250)distorted all statistics
- Lognormal distribution yields unrealistically high values
- Data are gamma distributed
  - Use of UPLs or UTLs based upon gamma distribution is suggested

Method	With outliers n=17	Without outliers n=15
Normal UTL95/95	3859	1780
Gamma UTL95/95 (WH)	5595	3015
Lognormal UTL95/95	18203	9666
Nonparametric UTL95/95	4300	1490
Gamma 95% Percentile	3249	1711
Lognormal 95% Percentile	4869	2681
Maximum	4300	1490

Elevated values

LOCKHEED MARTIN

Steps to Estimate BTVs

- Make sure no significant outliers or multiple populations are present in background data set
  - Use graphical displays to visualize data
  - Graphical methods provide useful information about outliers, multiple populations
- Perform GOF test to determine data distribution
  - Depending upon data distribution, use an appropriate parametric or nonparametric estimate of BTV
  - May want to consult a statistician for further clarification

LOCKHEED MARTIN

ProUCL 4.1.00

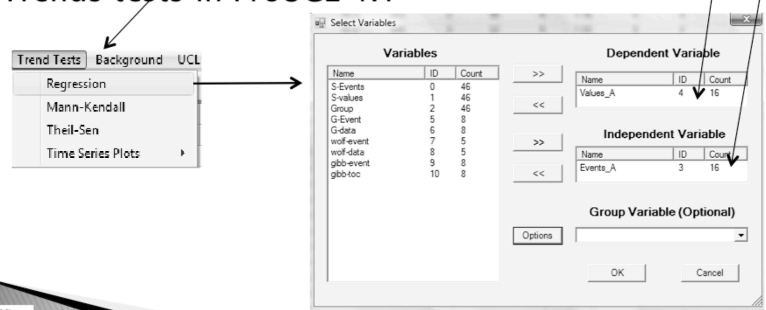
Trend Analysis: Linear Regression, Mann-Kendall Trend Test, and Theil-Sen Trend Line

http://www.epa.gov/osp/hstl/tsc/software.htm

62

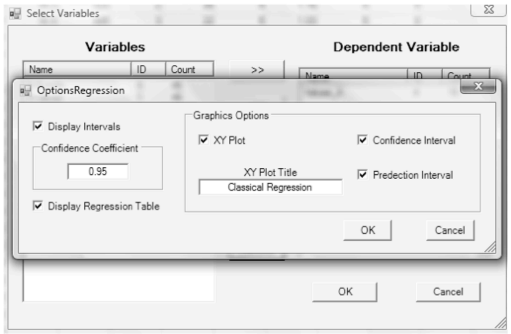
Trend Tests

- Trend tests are used to determine if concentrations (e.g., in a compliance well) are decreasing/increasing over time.
- A time variable and concentration variable with equal number of observations should be present in data set
- Trends tests in ProUCL 4.1



Linear Regression Line

- Linear regression Line Test



Slope of line determines trend in data

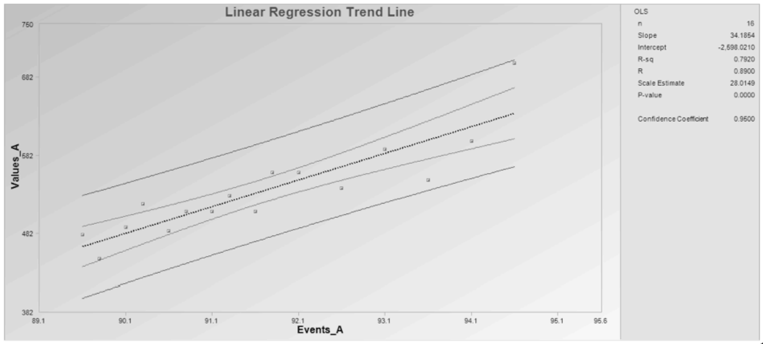
Significant positive slope suggests upward trend

Significant negative slope suggests downward trend

Insignificant slope suggests no evidence of trend in data

Significance is determined using p-value of slope test

Linear Regression Trend Line



For slope test p-value=0

Slope is significantly positive with p-value= 0.0

Conclude: there is significant upward trend in data

Regression Estimates and Inference Table

Paramater	Estimates	Std. Error	T-values	p-values
intercept	-2598	429.3	-6.051	2.9802E-5
Events_A	34.19	4.682	7.302	3.8968E-6

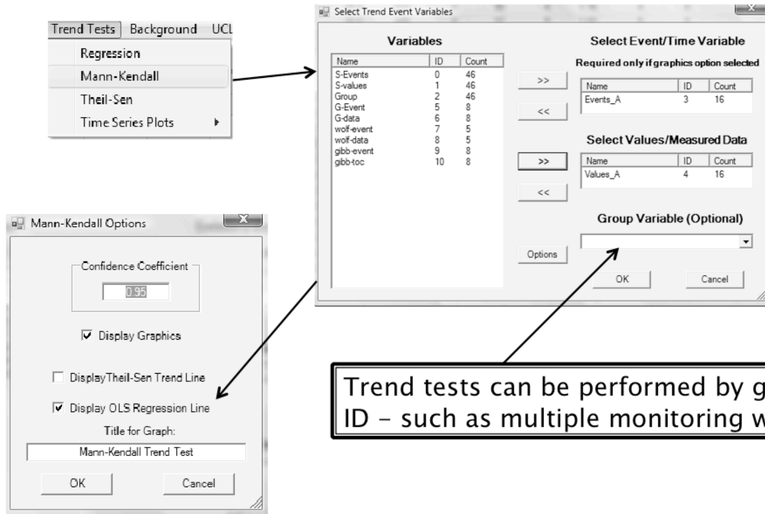
OLS ANOVA Table

Source of Variation	SS	DOF	MS	F-Value	P-Value
Regression	41848	1	41848	53.32	0.0000
Error	10988	14	784.8		
Total	52836	15			

R Square	0.792
Adjusted R Square	0.777
Sqrt(MSE) = Scale	28.01

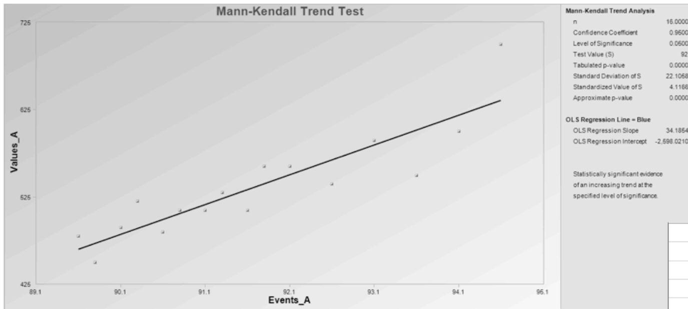
LOCKHEED MARTIN

Mann-Kendall Trend Test Statistic, S



Trend tests can be performed by group ID – such as multiple monitoring wells

Mann-Kendall Trend Test Statistic, S



S Test Statistic =92 >0

General Statistics	
Number of Events	16
Number of Values	16
Minimum	450
Maximum	700
Mean	536.6
Median	525
Standard Deviation	59.35
SEM	14.84
Mann-Kendall Test	
Test Value (S)	92
Tabulated p-value	0
Standard Deviation of S	22.11
Standardized Value of S	4.117
Approximate p-value	1.9228E-5
Statistically significant evidence of an increasing trend at the specified level of significance.	

Significant positive value of S or standardized S suggests upward trend

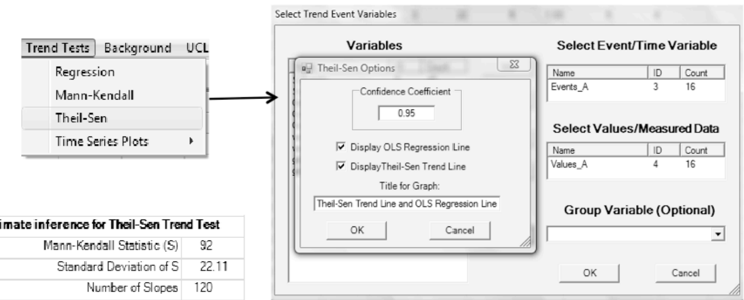
Significant negative value of S or standardized S suggests downward trend

Significance is determined using p-value

S >0 and p-value =0
Conclusion: Data exhibit upward trend

LOCKHEED MARTIN

Theil-Sen Trend Line Test



Approximate inference for Theil-Sen Trend Test	
Mann-Kendall Statistic (S)	92
Standard Deviation of S	22.11
Number of Slopes	120
Theil-Sen Slope	31.71
Theil-Sen Intercept	-2375
M1	38.34
M2	81.66
95% LCL of Slope (0.025)	20
95% UCL of Slope (0.975)	43.3
M1'	41.82
One-sided 95% lower limit of Slope	21.56
Statistically significant evidence of an increasing trend at the specified level of significance.	

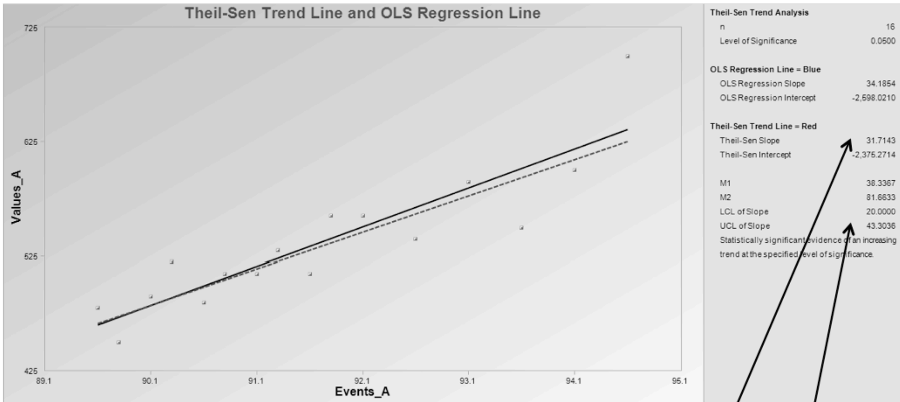
Theil-Sen Slope is positive

95% confidence interval (CI) for slope does not contain ‘0’.

This suggest that there is significant upward (positive) trend in data at 0.05 level of significance

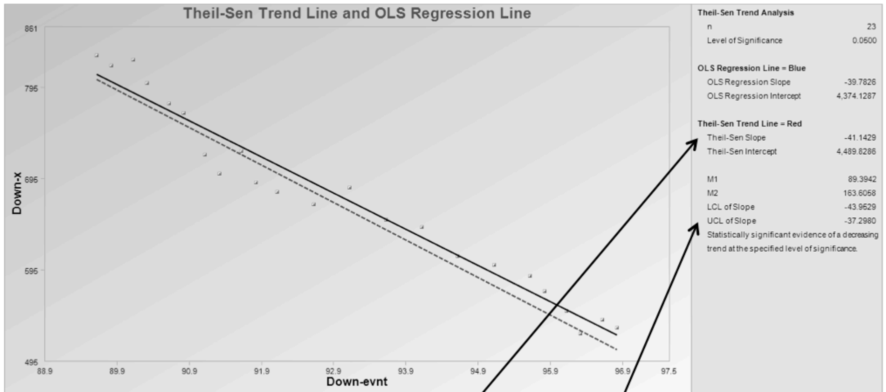
LOCKHEED MARTIN

Theil-Sen Trend Line Test – Upward Trend



Conclusion: Data exhibit significant upward trend at 0.05 level

Theil-Sen Trend Line Test-Downward Trend



Theil-Sen slope is negative

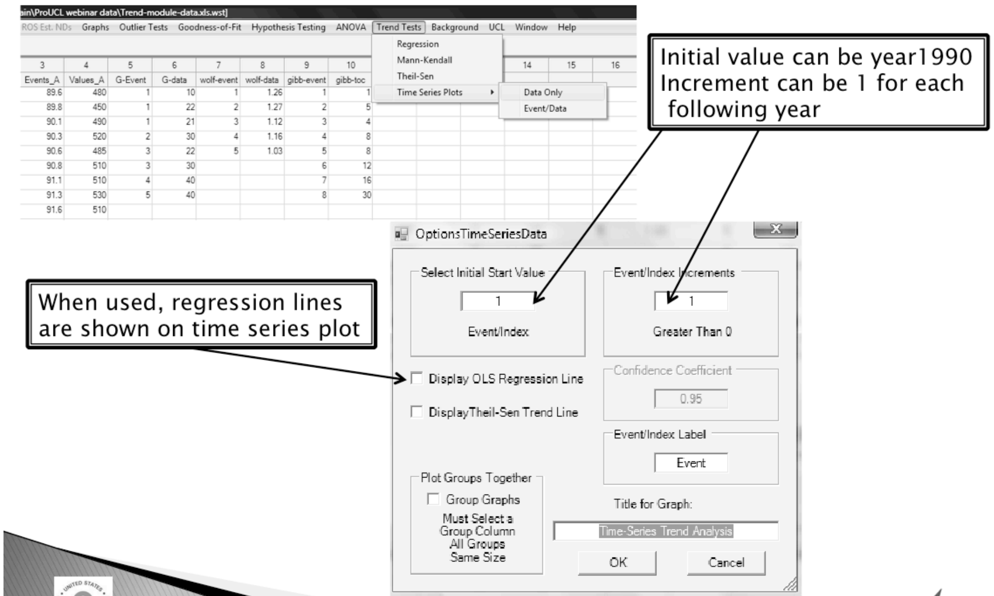
LCL and UCL of 95% CI of slope are <0 implying '0' is not in 95% CI

Conclusion: Data exhibit significant downward trend at 0.05 level

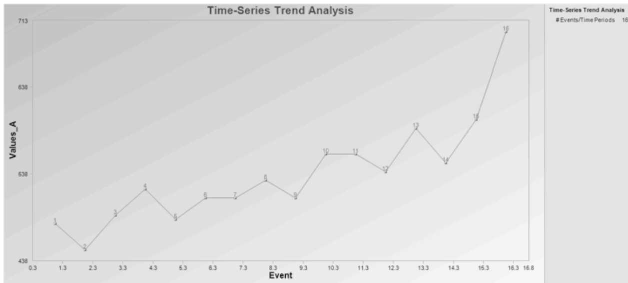
If CI for slope contains '0' – data do not exhibit significant trend

LOCKHEED MARTIN

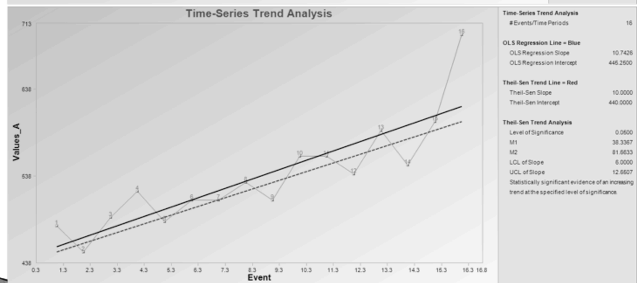
Time Series Plots with Data Only Option



Time Series Plots



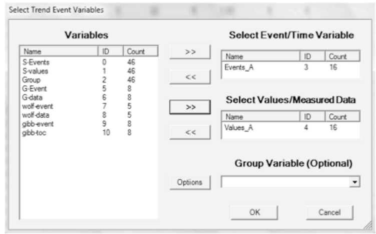
Simple index plot with chosen time events



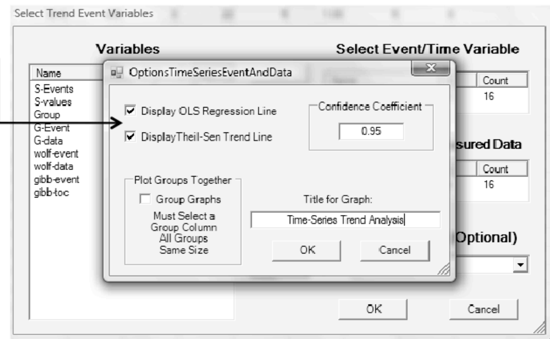
Plot as a function of chosen time events with trend lines

LOCKHEED MARTIN

Time Series Plot for Data vs. Sampling Events

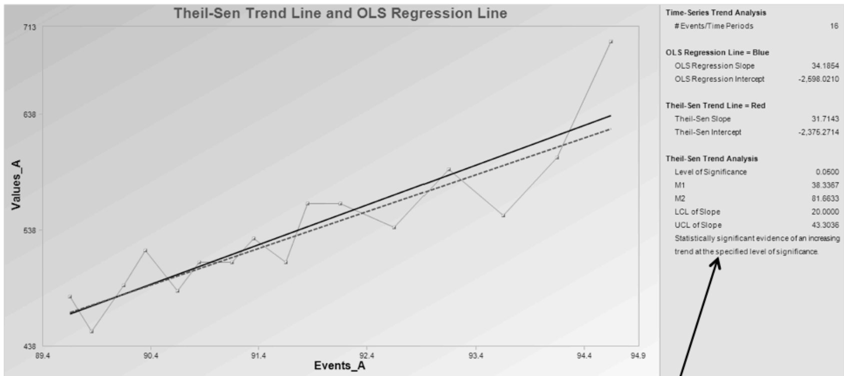


When checked regression lines are displayed on time series plot



Time Series Plot Identifying Trend in Data

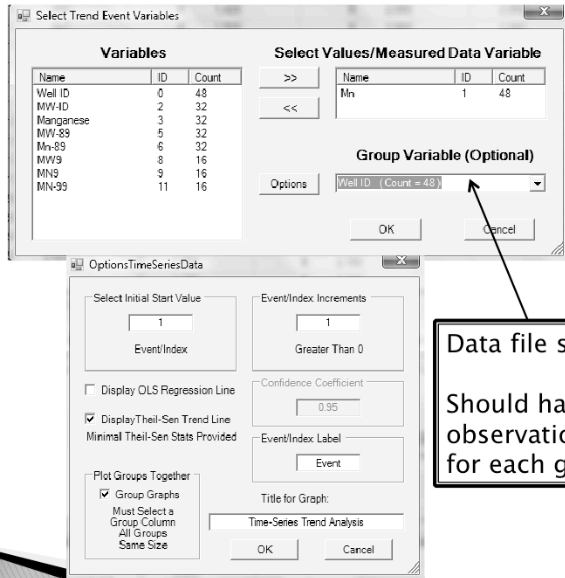
- Time Series plot identifying trend as a function of events



- Graph exhibits upward trend which is confirmed by trend test statistics

LOCKHEED MARTIN

Time Series Plots – Comparing Concentrations of Multiple Groups (Wells)



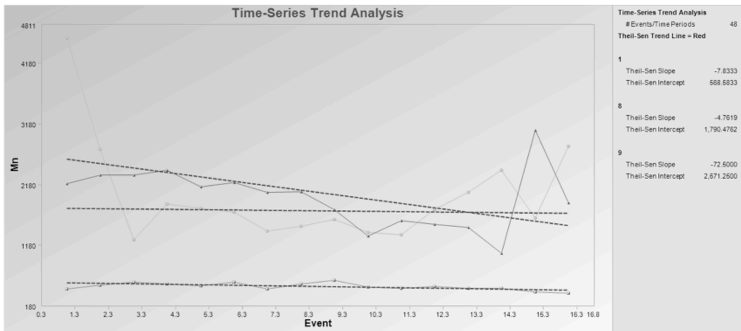
Data file should have a group ID

Should have same number of observations/ sampling events for each group (e.g., MW)

LOCKHEED MARTIN

Time Series Plots – Comparing Arsenic in Upgradient and Monitoring Wells

- Groundwater data from 3 MW wells: Well 1 is upgradient well, and wells 8 and 9 are MW wells



- Graph suggests that As in MW 8 and MW 9 are much higher than upgradient well 1.

LOCKHEED MARTIN

ProUCL 4.1.00

Computing Mean, Variance, UCL95, UPL95, and Upper Tolerance Limit for Data Sets with Nondetect (ND) Observations

http://www.epa.gov/osp/hstl/tsc/software.htm

77

Estimation Methods for Data Sets with NDs

- Parametric methods:
  - MLE (Normal), Regression on order statistics (ROS)
- Nonparametric methods:
  - Substitution, Kaplan-Meier (KM), and bootstrap methods
- For data sets with NDs, nonparametric methods (e.g., KM method) are preferred as distributional assumptions are hard to justify, especially when
  - NDs exceed detected values
  - Multiple detection limits (DLs) are present

LOCKHEED MARTIN

Upper Limits for Data Sets with NDs

- TCE data of size 14 has 6 NDs with a single DL = 0.68
  Data are :0.81 <0.68 <0.68 0.95< 0.68 <0.68 < 0.68 <0.68 9.29 1.9 0.88 2.98 0.75 5.97

Data File

0	1
TCE	D_TCE
0.81	1
0.68	0
0.68	0
0.95	1
0.68	0
0.68	0
0.68	0
0.68	0
9.29	1
1.9	1

Summary statistics using detects only

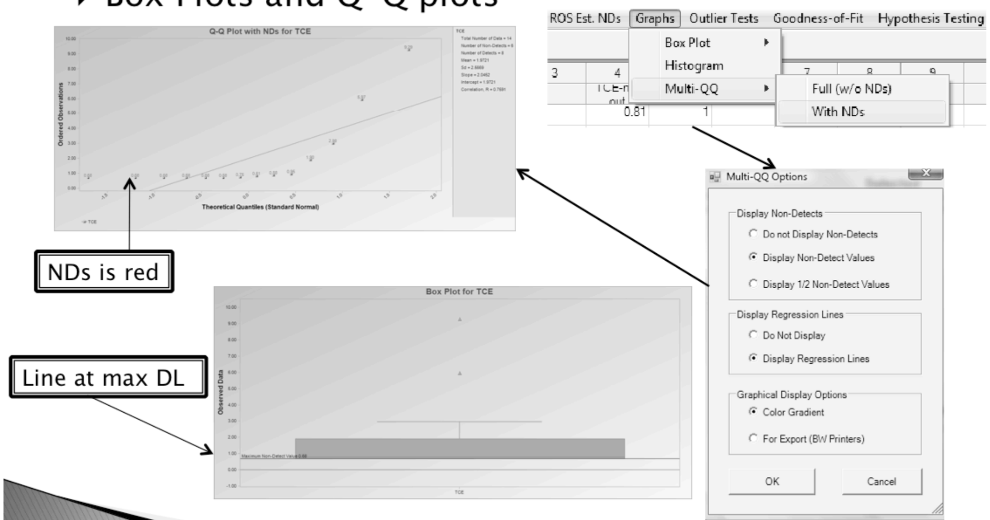
Summary Statistics for Raw Data Sets with NDs using Detected Data Only

				Raw Statistics using Detected Observations							
Variable	Num Ds	NumNDs	% NDs	Minimum	Maximum	Mean	Median	SD	MAD/0.675	Skewness	CV
TCE	8	6	42.86%	0.75	9.29	2.941	1.425	3.12	0.956	1.541	1.061
TCE-no-out	7	6	46.15%	0.75	5.97	2.034	0.95	1.917	0.297	1.822	0.943

Data from Drs. Warren and Nussbaum's Workshop at 2010 NARPM conference

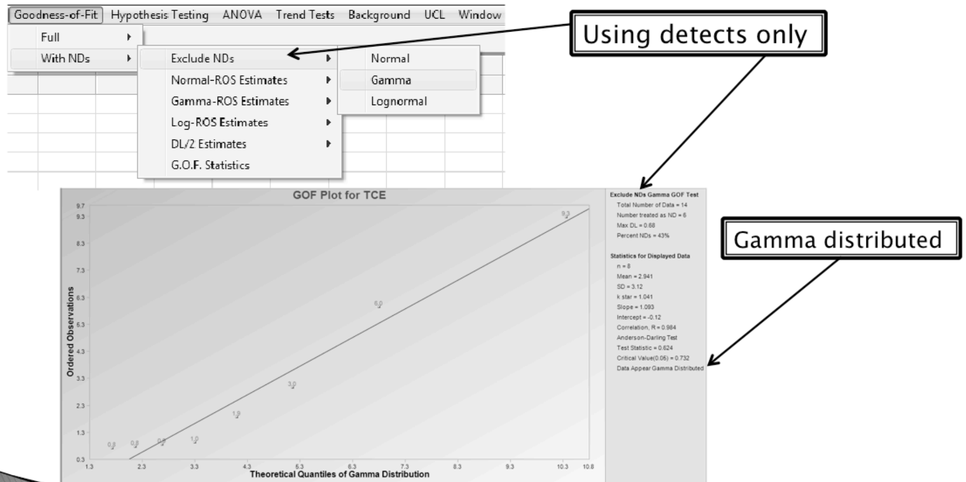
TCE Data Set with NDs

- Box Plots and Q-Q plots



Gamma GOF Test on Detected data

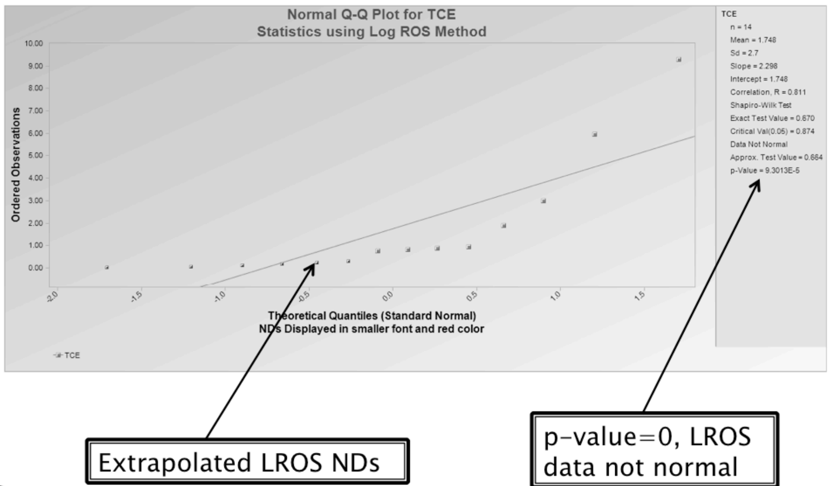
- Detected data appear to follow Gamma distribution



LOCKHEED MARTIN

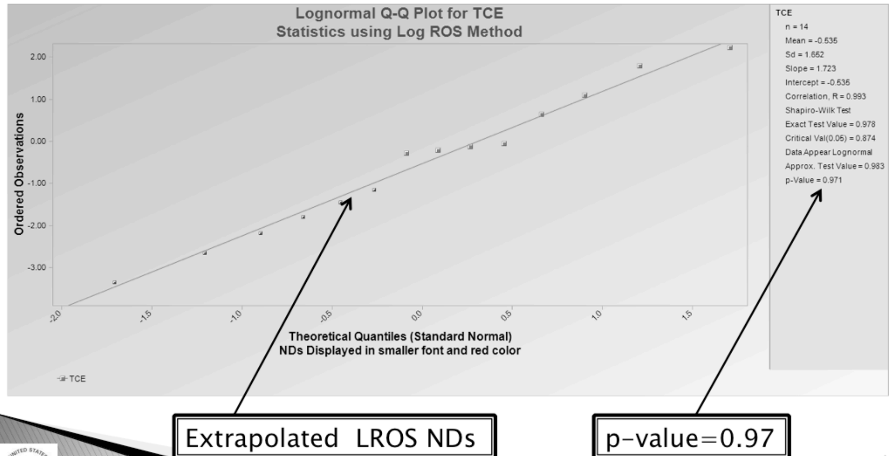
Lognormal ROS Method

- LROS data not normal



Lognormal ROS Method

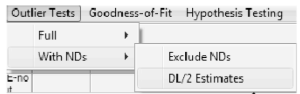
- LROS data are lognormal; Use parametric H-UCL or bootstrap method. Data set is of small size, H-UCL can be very large
- If nonparametric bootstrap methods used – just use KM method



LOCKHEED MARTIN

TCE Data Set with NDs

- Any outliers?



- At least one outlier = 9.29 (Dixon test), NDs=DL/2

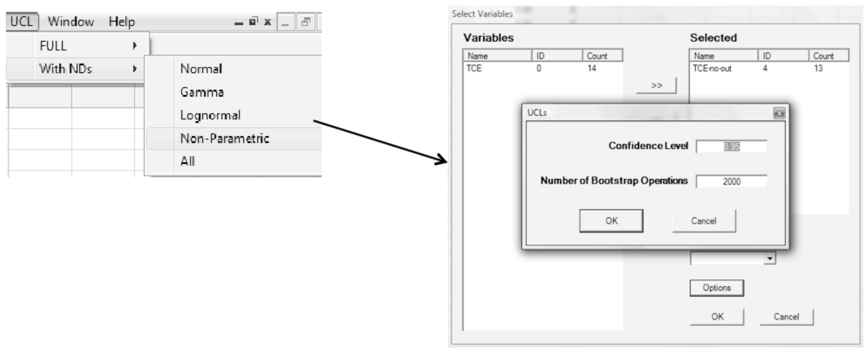
Dixon's Outlier Test for TCE		
Mean	1.972	
Standard Deviation	2.567	
Number of data = 14		
10% critical value: 0.492		
5% critical value: 0.546		
1% critical value: 0.641		
1. 9.29 is a Potential Outlier (Upper Tail)		
Test Statistic: 0.733		
For 10% significance level, 9.29 is an outlier.		
For 5% significance level, 9.29 is an outlier.		
For 1% significance level, 9.29 is an outlier.		

Project team should decide about disposition of 9.29

-Use it or not use it in computations

LOCKHEED MARTIN

UCLs for TCE Data with NDs



UCLs for TCE Data with NDs (with 9.29) - LROS

DL/2 Substitution Method			
Mean	1.826	95% H-Stat (DL/2) UCL	4.462
SD	2.65		
Mean (in Log Scale)	-0.1		
SD (in Log Scale)	1.135		
Log ROS Method			
Mean	1.748	95% t UCL	3.026
SD	2.7	95% Percentile Bootstrap UCL	3.039
Mean (in Log Scale)	-0.535	95% BCA Bootstrap UCL	3.442
SD (in Log Scale)	1.652	95% H-UCL (Log ROS)	14.23
Kaplan Meier (KM) Method			
Mean	2.002	95% KM (t) UCL	3.246
SD	2.458	95% KM (BCA) UCL	3.433
SE of Mean	0.702	95% KM (% Bootstrap) UCL	3.239
		95% KM (Chebyshev) UCL	5.063
		97.5% KM (Chebyshev) UCL	6.388
		99% KM (Chebyshev) UCL	8.99
Potential UCL to Use		Data appear Gamma Distributed (0.05)	
		May want to try Gamma UCLs	

LROS Stats

KM Stats

UCLs for TCE Data with NDs (with 9.29) –GROS

95% UCLs (Adjusted for Skewness)	
95% Adjusted-CLT UCL (Chen-1995)	3.333
95% Modified-t UCL (Johnson-1978)	3.049
95% Nonparametric UCLs	
95% Bootstrap-t UCL	4.905
95% Hall's Bootstrap UCL	7.983
95% Gamma UCLs(Assuming Gamma Distribution)	
95% Approximate Gamma UCL	6.843
95% Adjusted Gamma UCL	8.405

Stats obtained by replacing GROS extrapolated negative values by 0.0001

95% UCLs (Adjusted for Skewness)	
95% Adjusted-CLT UCL (Chen-1995)	3.333
95% Modified-t UCL (Johnson-1978)	3.049
95% Nonparametric UCLs	
95% Bootstrap-t UCL	5.231
95% Hall's Bootstrap UCL	8.015
95% Gamma UCLs(Assuming Gamma Distribution)	
95% Approximate Gamma UCL	8.859
95% Adjusted Gamma UCL	11.3

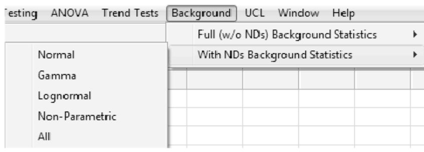
Stats obtained by replacing GROS extrapolated negative values by 0.000001

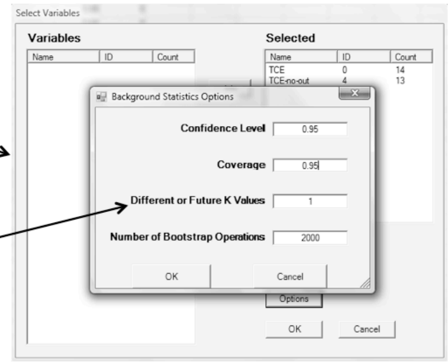
Caution: GROS yields negative extrapolated values which are replaced by small numbers: 0.001, 0.0001, …

Different choices for negative values yield different UCL95

Use of ROS methods should be avoided, use nonparametric KM method

Background Statistics – UPL, UTL for TCE Data with NDs





Prediction intervals for k≥1

UPL95, UTL95-95 TCE Data with NDs

- Statistics computed with outlier 9.29

Background Statistics Assuming Normal Distribution	
DL/2 Substitution Method	
Mean	1.826
SD	2.65
Tolerance Factor K	2.614
95% UTL 95% Coverage	8.755
95% UPL (t)	6.685
90% Percentile (z)	5.223
95% Percentile (z)	6.186
99% Percentile (z)	7.992
Note: DL/2 is not a recommended method.	
Maximum Likelihood Estimate (MLE) Method	
Mean	0.72
SD	3.679
95% UTL 95% Coverage	10.34
95% UPL (t)	7.465
90% Percentile (z)	5.436
95% Percentile (z)	6.772
99% Percentile (z)	9.28

Log ROS Method	
Mean in Log Scale	-0.535
SD in Log Scale	1.652
Mean in Orignal Scale	1.748
SD in Original Scale	2.7
95% UTL 95% Coverage	43.98
95% BCA UTL with 95% Coverage	9.29
95% Bootstrap (%) UTL with 95% Coverage	9.29
95% UPL (t)	12.1
90% Percentile (z)	4.866
95% Percentile (z)	8.869
99% Percentile (z)	27.34
Kaplan-Meier (KM) Method	
Mean	2.002
SD	2.458
Standard Error of Mean	0.702
95% UTL 95% Coverage	8.427
95% KM Chebyshev UPL	13.09
95% KM UPL (t)	6.508
90% KM Percentile (z)	5.152
95% KM Percentile (z)	6.045
99% KM Percentile (z)	7.72

BTVs using LROS

BTVs based upon KM Method

ProUCL 4.1.00

Case Study
Estimating Background Threshold Values (BTVs)
Arsenic Data with Nondetects from a Federal Facility

90

Arsenic Data Set from a Federal Facility

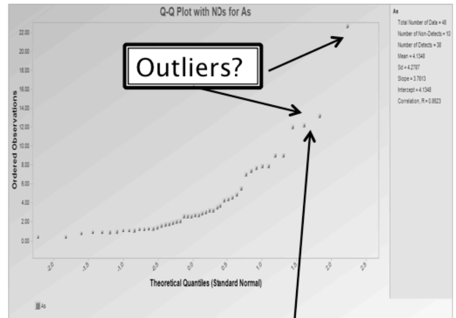
Objective: Compute site-specific background level concentrations/BTVs to compare site data with BTVs

- Look at your background/reference data graphically
- Perform outlier tests
  - Establish background/reference area data set represented by unimpacted locations
  - Perform GOF test to determine data distribution
- Compute UPLs and UTLs to estimate BTV
  - For added insight – use formal graphical comparisons of AOC data with reference BTVs

91

LOCKHEED MARTIN

Box Plot and Q-Q Plot of Arsenic Background Data Set of Size 48



Box Plot of Arsenic	Q-Q Plot of Arsenic – NDs shown in red

Graphs suggest 22.7 is an outlier; and potentially a group of 3 intermediate outliers (?) is also present

92

Rosner Outlier Test on Background Data

Rosner's Outlier Test for As

Mean	**4.135**
Standard Deviation	**4.279**
Number of data	**48**
Number of suspected outliers	**5**

#	Mean	sd	Potential outlier	Obs. Number	Test value	Critical value (5%)	Critical value (1%)
1	4.135	4.234	22.7	19	4.385	3.11	3.46
2	3.74	3.325	13.2	18	2.845	3.1	3.46
3	3.534	3.044	12.2	21	2.847	3.09	3.45
4	3.342	2.781	12	14	3.114	3.09	3.44
5	3.145	2.476	9	7	2.365	3.08	3.43

For 5% significance level, there are 4 Potential Outliers
Therefore, Potential Statistical Outliers are
22.7, 13.2, 12.2, 12

For 1% Significance Level, there is 1 Potential Outlier
Therefore, Observation 22.7 is a Potential Statistical Outlier

22.7, 13.2, 12.2, 12 are statistical outliers at 0.05 level of significance

4 outliers:
Test val. > critical val.
Conclude: 4 outliers

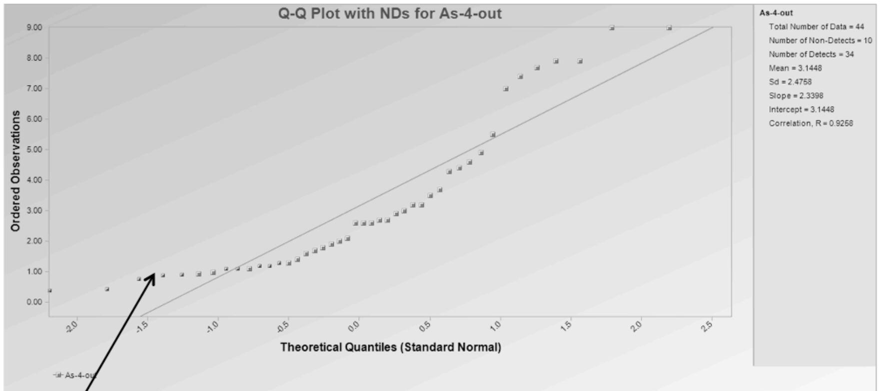
5 outliers:
Test val. < critical val.
Conclude: not 5 outliers

Project team should decide about disposition of outliers

Supplement outlier tests with graphical displays

LOCKHEED MARTIN

Q-Q Plot of Background As Without 4 Outliers



Q-Q plot of Arsenic without 4 outliers - NDs shown in red
Graph does not display a linear pattern
Background arsenic does not follow normal distribution

LOCKHEED MARTIN

Summary Statistics for Background/ Reference Area and AOC Data Sets

Site AOC Data

Background Data

From File: C:\Users\Anita\Desktop\ProUCL-webinar-68\webinar-main\proucl webinarslides2\SS_AOC_BK_NO_As-use.xls.wst

Summary Statistics for Raw Data Sets with NDs using Detected Data Only

				Raw Statistics using Detected Observations							
Variable	Num Ds	NumNDs	% NDs	Minimum	Maximum	Mean	Median	SD	MAD/0.675	Skewness	CV
As_out4 (a1)	66	13	16.46%	0.78	10.7	3.718	3.45	1.75	1.26	1.405	0.471
As_out4 (a14)	97	0	0.00%	0.42	7	3.293	3.3	1.613	1.631	0.226	0.49
As_out4 (a22)	22	2	8.33%	2.1	4.5	3.427	3.4	0.672	0.593	-0.222	0.196
As_out4 (aoc-10)	19	0	0.00%	0.65	6	3.92	4.2	1.502	1.334	-0.851	0.383
As_out4 (aoc-11)	68	0	0.00%	0.73	87.9	6.471	3.8	11.27	1.853	6.042	1.741
As_out4 (aoc-12)	53	0	0.00%	0.34	92.1	7.472	2.59	18.12	1.468	3.892	2.425
As_out4 (aoc-20)	42	9	17.65%	0.36	15.1	2.747	2.345	2.388	1.29	3.669	0.869
As_out4 (aoc-6)	10	0	0.00%	2.5	5.7	3.63	3.45	1.034	0.815	1.059	0.285
As_out4 (aoc-7)	14	0	0.00%	1.2	27.2	8.493	3.5	9.115	2.817	1.132	1.073
As_out4 (bk)	34	10	22.73%	0.93	9	3.797	2.95	2.456	2.001	0.905	0.647
As_out4 (ge)	19	3	13.64%	0.83	5.6	3.107	3.3	1.402	1.779	0.151	0.451
As_out4 (s2)	115	4	3.36%	1	59.9	6.61	5.1	7.614	2.224	4.918	1.152
As_out4 (s5)	60	10	14.29%	0.63	6.5	2.733	2.35	1.504	1.631	0.526	0.55
As_out4 (sa1-tnt)	71	0	0.00%	0.51	144	7.765	3.5	18.99	2.52	5.882	2.446
As_out4 (sa4-hsp)	3	1	25.00%	0.78	2.2	1.693	2.1	0.793	0.148	-1.701	0.468

4 outliers excluded from background (As-out4 (bk)) data set

95

LOCKHEED MARTIN

GOF Test for Reference Data Set

- Background data set has 10 nondetects
- For data sets with many NDs, it is not easy (nor needed) to justify distributional assumptions
- Use of nonparametric method such as KM method is suggested to compute BTVs

KM (1958) method is a nonparametric method used on data sets with NDs to estimate population mean, standard deviation, standard error of mean, UCLs, UPLs, and UTLs

LOCKHEED MARTIN

Estimates of BTV Using KM Method

Background data set of size 48 was collected
Background data screened for outliers
Identified outliers removed – Project team should get involved

Kaplan-Meier (KM) Method	
Mean	3.147
SD	2.441
Standard Error of Mean	0.374
95% UTL 95% Coverage	8.251
95% KM Chebyshev UPL	13.91
95% KM UPL (t)	7.297
90% KM Percentile (z)	6.276
95% KM Percentile (z)	7.162
99% KM Percentile (z)	8.826
95% KM Simultaneous Upper Limit	10.24

BTV estimates without 4 outliers

Kaplan-Meier (KM) Method	
Mean	3.742
SD	3.284
Standard Error of Mean	0.486
95% UTL 95% Coverage	10.55
95% KM Chebyshev UPL	18.21
95% KM UPL (t)	9.313
90% KM Percentile (z)	7.951
95% KM Percentile (z)	9.144
99% KM Percentile (z)	11.38
95% KM Simultaneous Upper Limit	13.37

BTV estimates without 1 outlier

Expanding Site-Specific Background Data

- PRP suggested to expand existing background data by including onsite AOC data comparable to background
- Since a background data set of size 48 is already available, BTV is computed using background data; and
- Observations less than BTV are considered coming from background population (common practice in background evaluation studies) establishing the expanded background data set.

Extracting Site-Specific Background Data

- However, when background data are not available, one can potentially extract background data from onsite data using normal Q-Q plots and population partitioning methods (e.g., Singh, Singh, Flatman, 1994, Math Geology).

- Since background data were already collected, population partitioning was not used to extract background data from onsite data.

    - A scenario extracting background data using onsite data is considered for illustration purposes.

Expanding Site-Specific Background Data Based upon BTV

- All onsite arsenic values less than BTV = UTL95-95 (=8.25 without 4 outliers) are considered as coming from background population.
- This resulted in expanded background data of 639.
- Formal Index Plots and Q-Q plots of reference and AOC data sets using original background data (without 4 outliers and 1 outlier) and expanded background data are shown next.

100

LOCKHEED MARTIN

These graphs with horizontal lines at BTV estimates represent formal graphical displays.

Onsite observations exceeding the BTVs may represent locations not belonging to site background population.

These onsite locations may require further investigation.

Statistical Software Scout 1.1

- Both ProUCL4.1 and Scout1.1 software packages were used to establish expanded background data
- Scout with advanced graphical capabilities and robust statistical methods was developed by Lockheed Martin for NERL-EPA Las Vegas
- An older Scout 2008 Version 1.0, its User and Technical Guides can be downloaded from EPA ESD website:

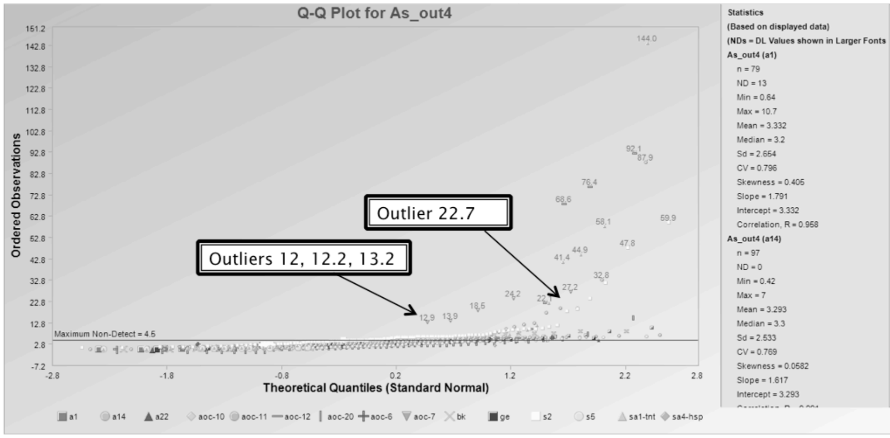
http://www.epa.gov/esd/databases/scout/abstract.htm

LOCKHEED MARTIN

Formal Graphical Displays Generated by Scout 1.1

- Scout can generate graphical displays including Index Plots and Q-Q plots by groups.
- These graphs can be *formalized* by displaying horizontal lines at decision statistics such as UPLs, upper percentiles, and UTLs.
  - On graphical displays, Scout can label observations by values, group ID, observation numbers- providing a formal visual comparison of background and AOC data sets.
- At present ProUCL cannot generate formal Index plots and cannot label observations by group IDs.

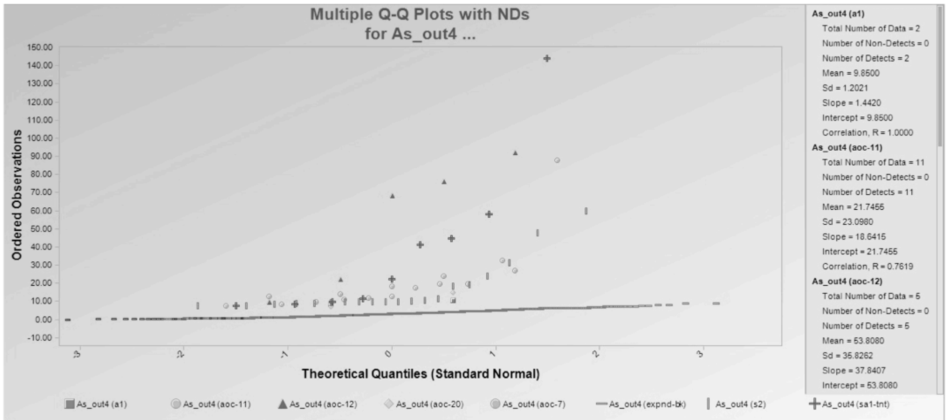
Q-Q plots Using Original Background (without 4 outliers) and AOC Data Sets



Graph compares As in AOCs and original background without 4 outliers (positions of removed outliers shown on graph).

LOCKHEED MARTIN

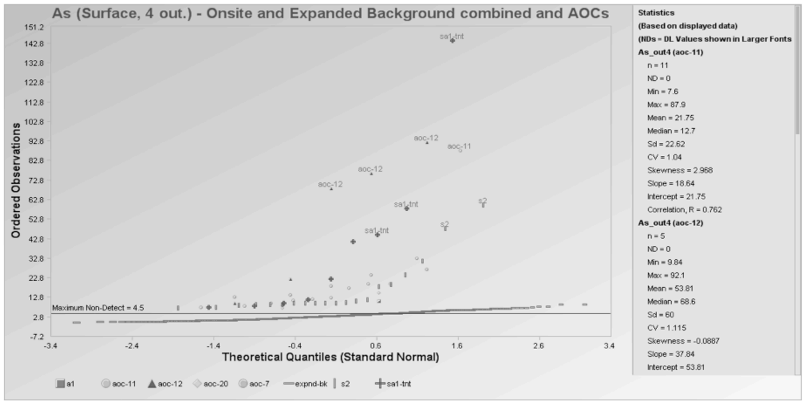
Q-Q Plots Using Expanded Background Data without 4 Outliers



Graph identifies onsite observations that are significantly higher than background population

Graph generated by ProUCL

Q-Q Plots Using Expanded Background Data without 4 Outliers

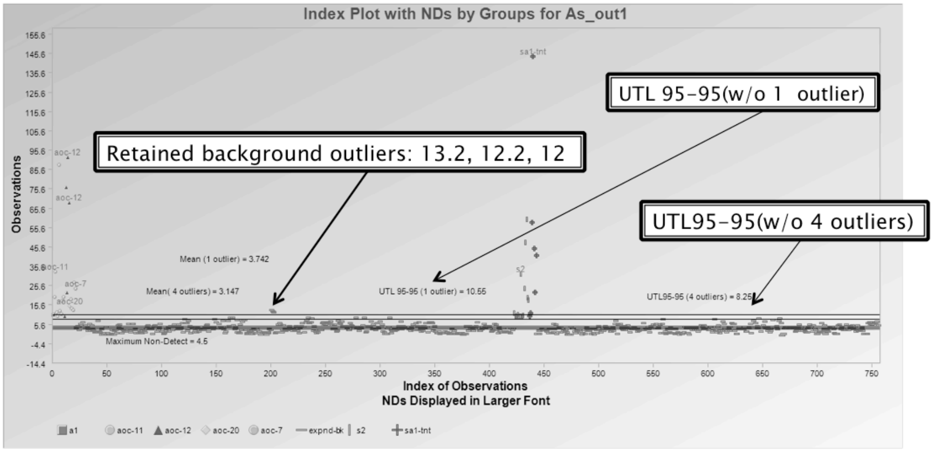


Graph identifies onsite observations that are significantly higher than background population

Graph generated by Scout, observations labeled by Group ID

LOCKHEED MARTIN

Formal Index Plot of Expanded Background Data (without 1 Outlier) and AOC Data Sets

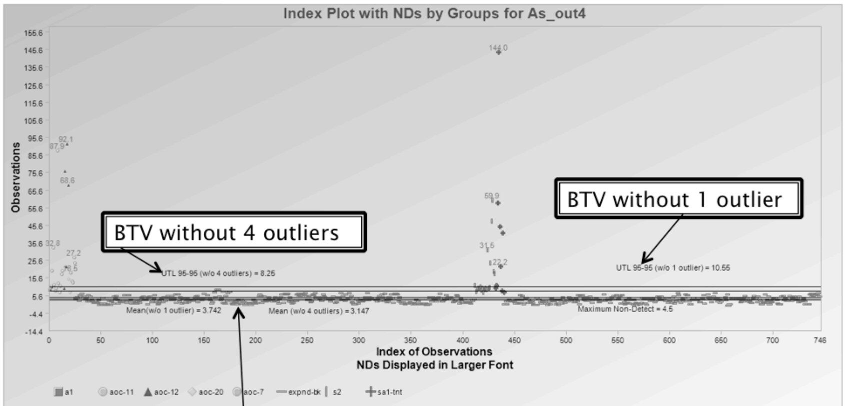


Graph shows BTVs (and mean) computed using background data without 1 and also without 4 outliers.

AOC-7, AOC-11, AOC-12, AOC-20, S2, Sa1-TNT exhibit higher AS than background

LOCKHEED MARTIN

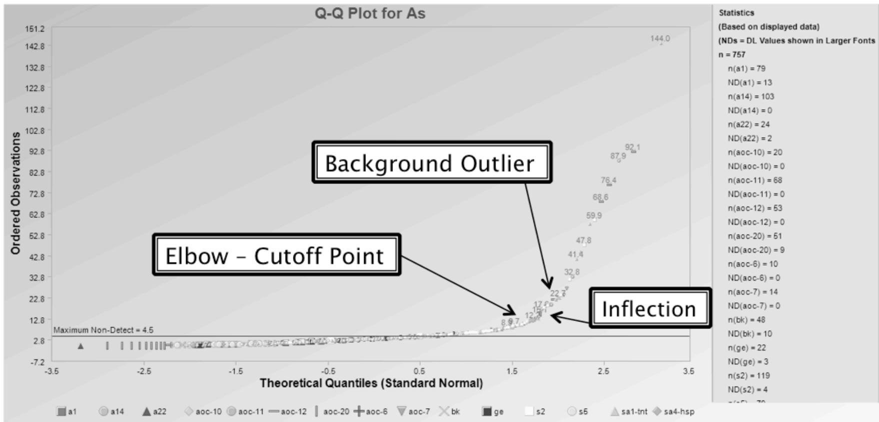
Formal Index Plot of Expanded Background Data (without 4 outliers) and AOC Data Sets



Observations above BTV may not belong to background
Mean w/o 1 outlier = 3.742, mean w/o 4 outliers = 3.147; both are below largest ND
AOC-7, AOC-11, AOC-12, AOC-20, S2, Sa1-TNT exhibit higher arsenic than BTVs

LOCKHEED MARTIN

Extracting Background Data from Onsite Data Using Population Partitioning



Q-Q plot of Arsenic using all data (757 points) from background and AOCs

LOCKHEED MARTIN

Extracting Background Data from Onsite Data Using Population Partitioning

- Background outlier 22.7 represents an impacted location

- Elbow of Q-Q plot seems to be around 9-13 and inflection point is around 15 (cutoff between Arsenic in background and onsite locations)
  - Project team should make a decision about which value to select as the cutoff point

- AOC arsenic less than 9-13 may be considered as representing site-specific background:
  - Onsite values <13 may be used to establish extracted site-specific background data and compute BTV estimates based upon extracted background data

109

LOCKHEED MARTIN

Summary – Establishing Background Data and Estimating BTVs

- Establishing background data sets:
- Collect appropriate amount of data from background locations
  - Make sure that no outliers and/or multiple populations are present in background data set
  - Use Q-Q plots to visualize data for additional insight about extremeness of outliers and/or multiple populations
    - Q-Q plots help Project Team in determining which values represent potential outliers not belonging to background population
    - This information cannot be obtained by using outlier tests alone

LOCKHEED MARTIN

Summary – Establishing Background Data and Estimating BTVs

- Based upon CSM/historical information, Project Team should decide about disposition of identified outliers – include or not include them in BTV calculations

- Perform GOF tests to determine data distribution

- Depending upon data distribution and frequency of nondetects in background data, use an appropriate (e.g., UTL95-95) parametric or nonparametric method to estimate of BTVs

  - Compute UTL95-95 to establish BTVs

LOCKHEED MARTIN

Summary – Establishing Background Data and Estimating BTVs

- In case of uncertainties (e.g., at large Federal Facilities), background data can be established in more than one way:
  - Collect data from background locations;
  - Expand background data by including onsite data comparable to collected background data;
  - Extract site-specific background data from available onsite data
    - Population partitioning methods based upon Q-Q plots can be used to extract site-specific background data from onsite data

LOCKHEED MARTIN

Summary – Establishing Background Data and Estimating BTVs

- Project Team should determine the appropriateness of extracted background data from onsite data

- Use of more than one method to establish background data set provides managers more options which can help them in:

  - Making cost effective and defensible decisions; and
  - Risk management evaluations

- Consult a statistician for clarification and discussing the best approaches to establish background data and estimate BTVs

Resources & Feedback

- To view a complete list of resources for this seminar, please visit the **Additional Resources**
- Please complete the **Feedback Form** to help ensure events like this are offered in the future

Need confirmation of your participation today?

Fill out the feedback form and check box for confirmation email.

114