PCC Statistics for Leachate Trend Analysis: Kendall's Tau

Kansas Department of Health and Environment

SWANA Conference, Manhattan, KS; November 20, 2014

Our Mission: To protect and improve the health and environment of all Kansans.

Presentation Objectives

To describe and demonstrate a statistical methodology for validating landfill emission data which meet the PCC criteria for activity reduction and/or termination.

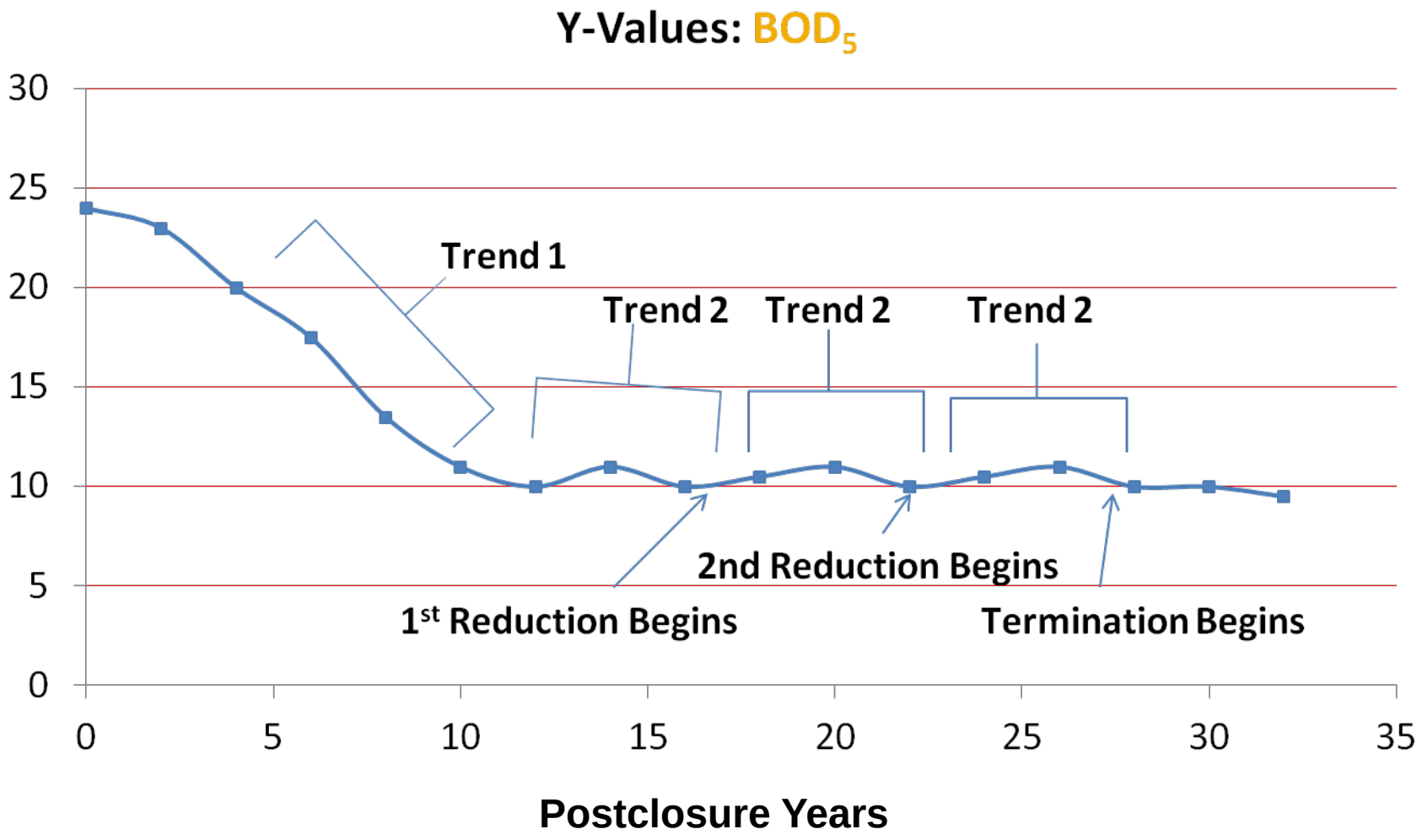
Specific PPP objectives include:

1. **Identify the PCC criteria.**
2. **Use Forest View leachate data to illustrate method.**
3. **Give an overview of statistics.**
4. **Describe and demonstrate the Kendall's tau method.**
5. **Identify available references to support the PPP.**

No. 1 - Identify the PCC criteria

- The BWM website provides policy and technical guidance (TGD) documents presented at the recent May 7th **PCC Reduction/Termination Workshop** in Salina.
- In particular, the **TGD SW-2014-G1** identifies the criteria for reducing and/or terminating PCC activities.
- The link to the specific criteria is: http://www.kdheks.gov/waste/techguide/SW-2013-G3.pdf
- The criteria are demonstrated in the next slide.

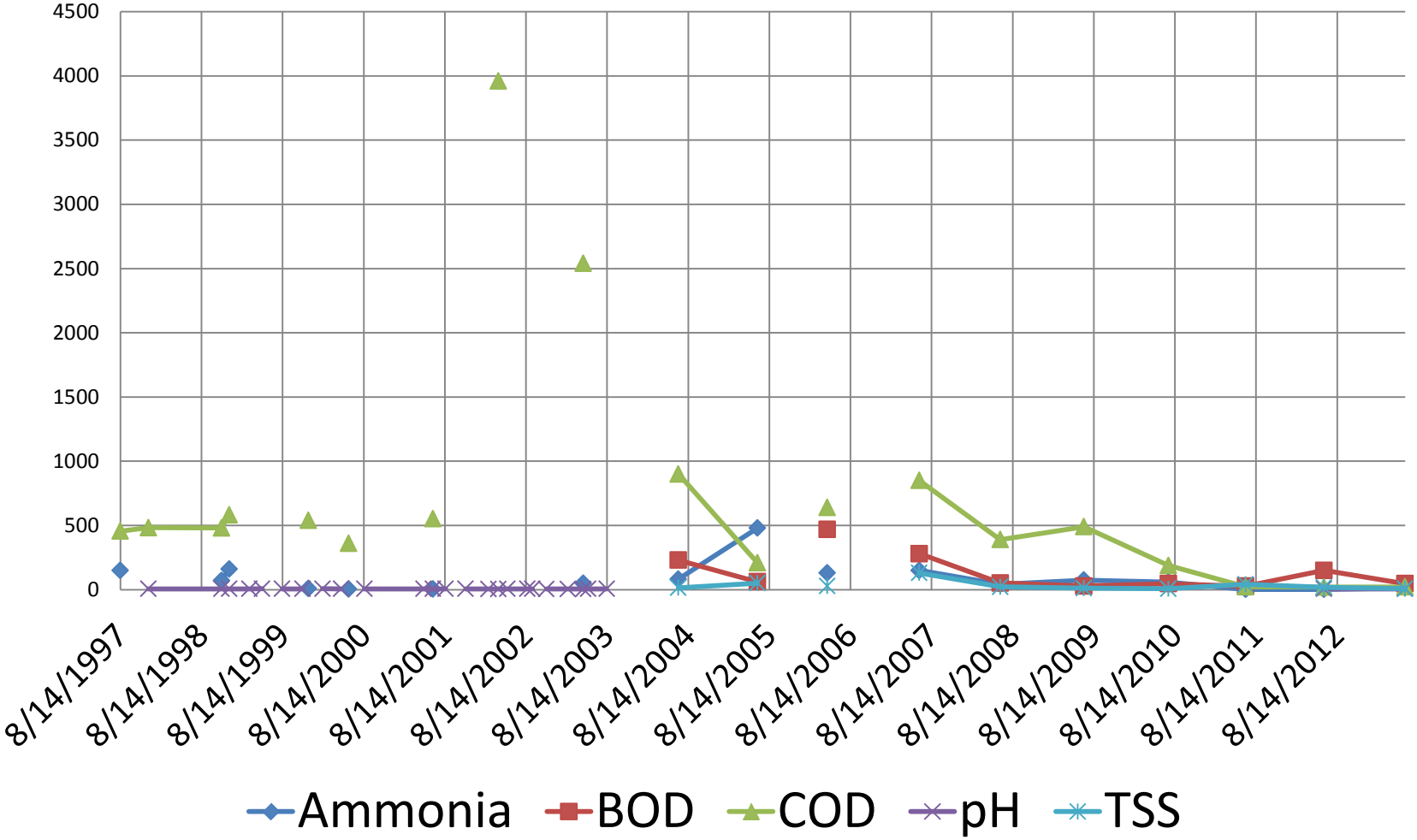
Figure 1 - Example 5 Year Trends for One Key Parameter:



No. 2 – Selection of Forest View

- It is one of two closed Subtitle D landfills in Kansas; the other is **Wheatland**.
- It has the most complete set of PCC leachate data of the two landfills.
- It stopped receiving MSW on December 29, 2006 and PCC began on December 2, 2010.
- Available leachate data started in August 1997 (see the following time series plot)

Figure 1 - Forest View Time Series Plot



No. 3 – Overview of Statistics

Merriam-Webster gives the following **definition of statistics**: "A branch of **mathematics** dealing with the collection, analysis, interpretation, and presentation of masses of numerical data."

There are **two types of statistics**: normal (or parametric) and non-normal (non-parametric). Most environmental data are of the latter type.

How to determine which type?

- You can plot a **histogram (ITRC Guidance)**:

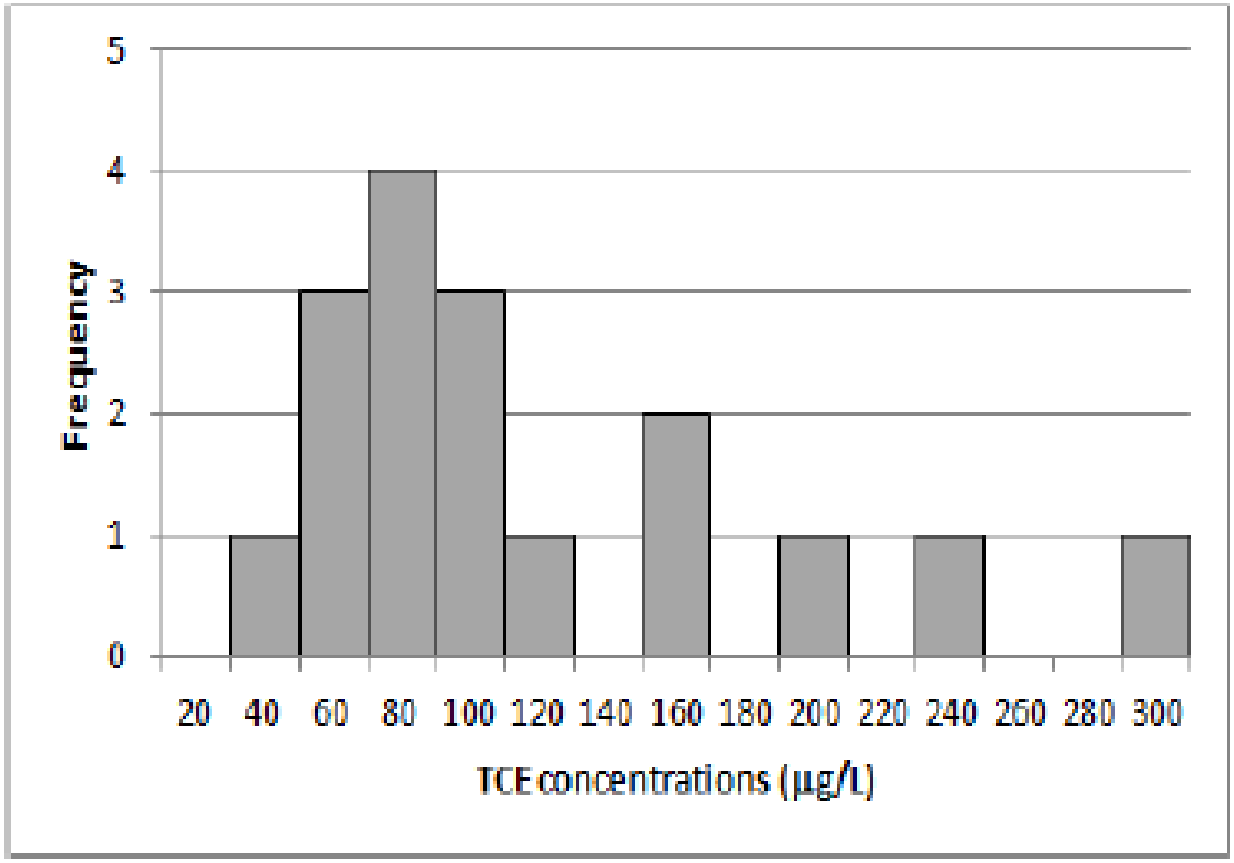


Figure 5-10. Data set as a histogram.

You can plot a **probability curve (ITRC Guidance)**

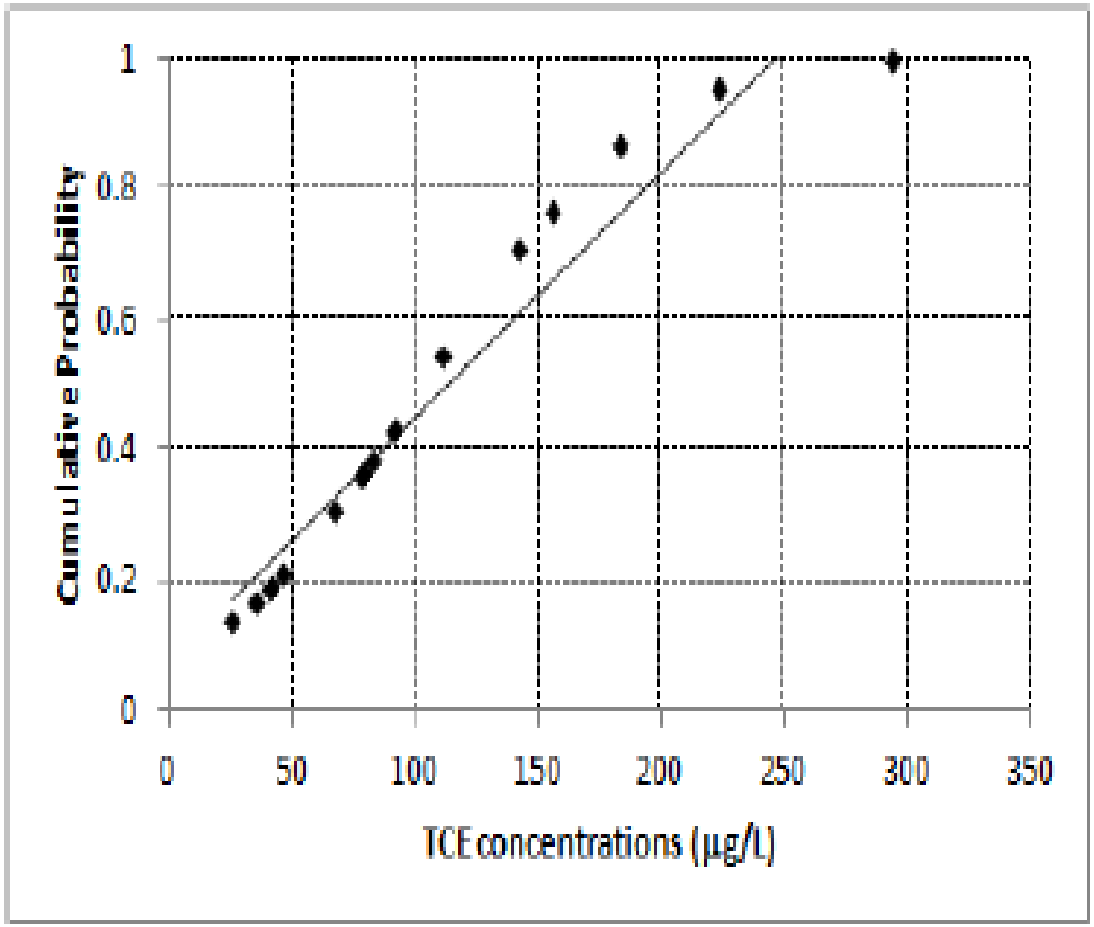


Figure 5-9. Data set as a probability plot.

Here's where the math comes in.

If the data plots as a **"bell shaped"** histogram, one could write an equation that would define the data. The same can be said if the probability plot is linear. Our sample data do not meet these criteria; hence, they are not normal or parametric, but non-normal or non-parametric. One can mathematically manipulate the data to give bell shaped and straight line curves as shown in the next plots.

Normalized Histogram (ITRC Guidance)

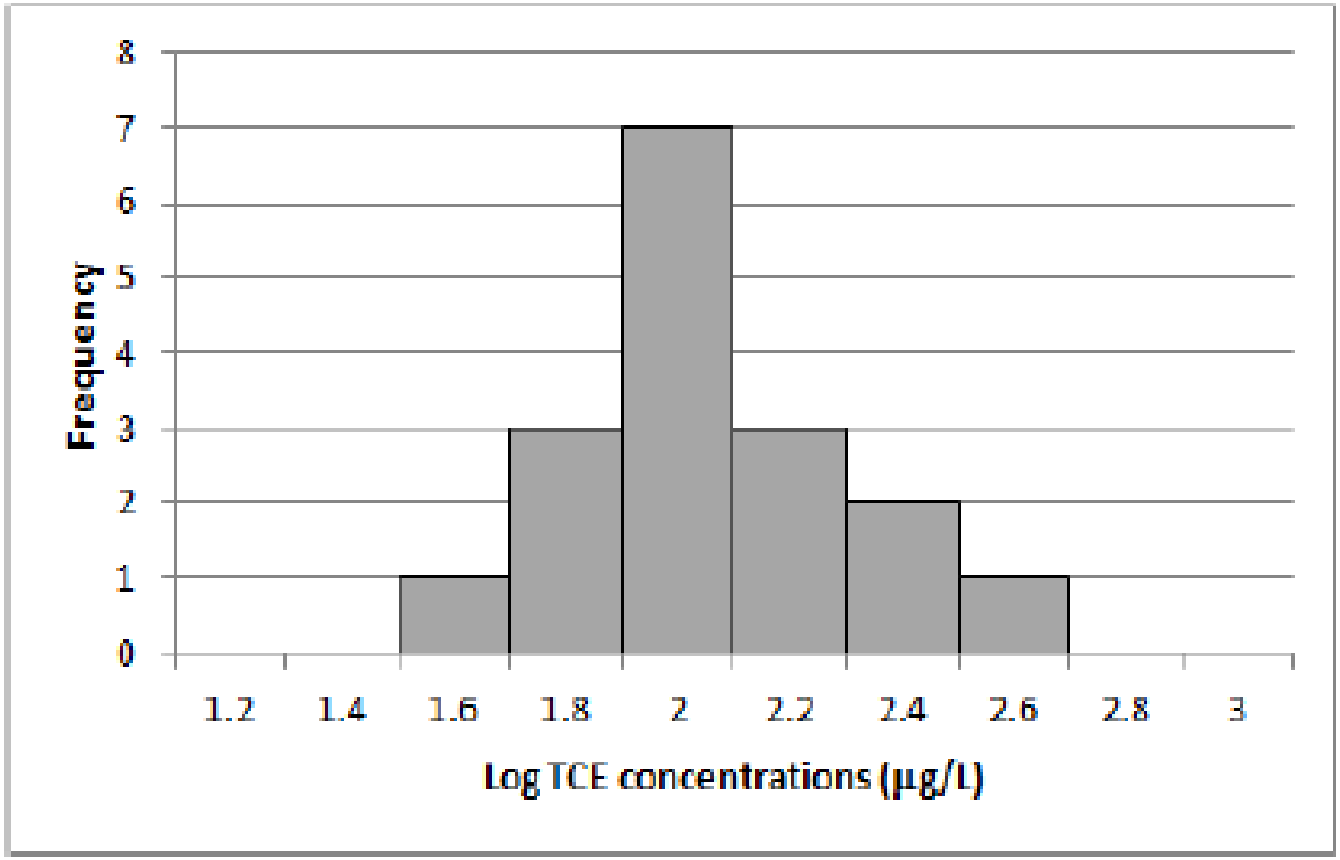


Figure 5-12. Histogram of log-transformed data.

Normalized Probability Curve (ITRC Guidance)

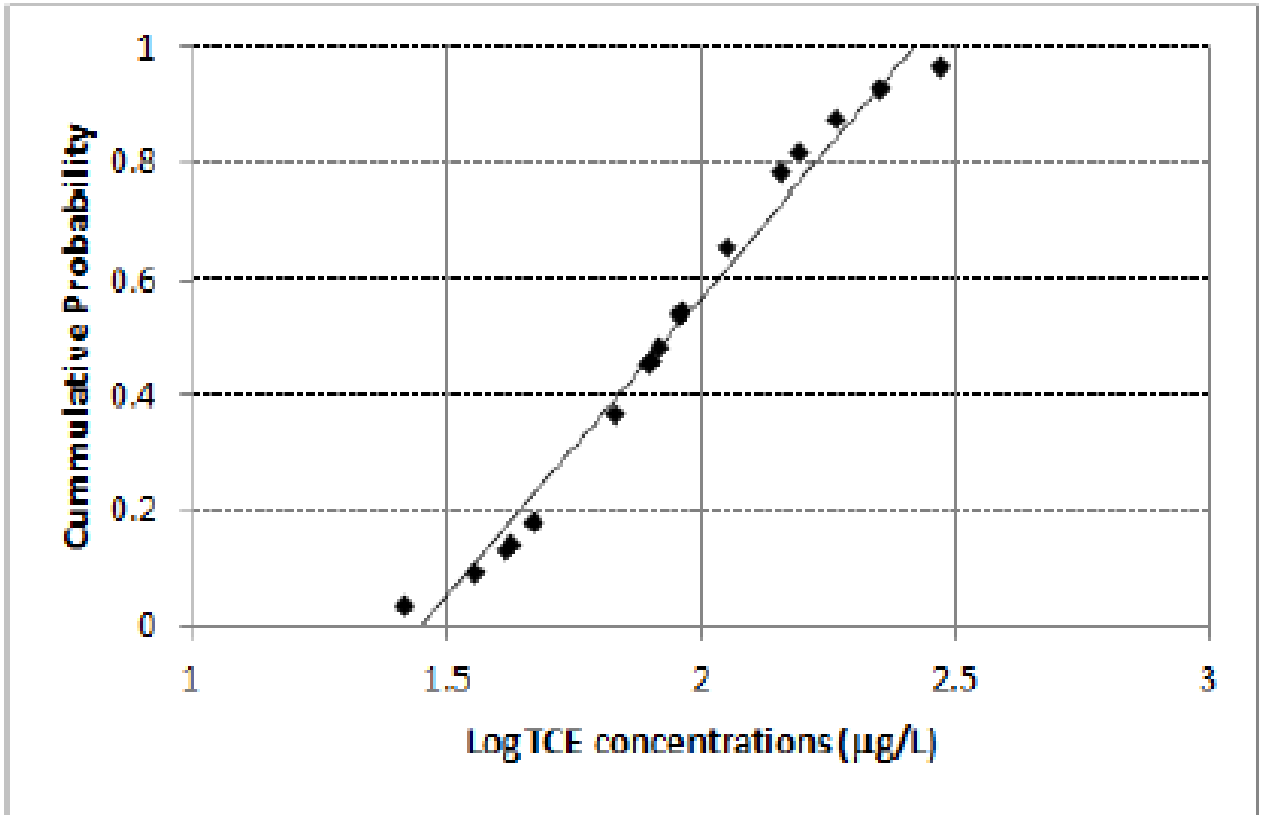
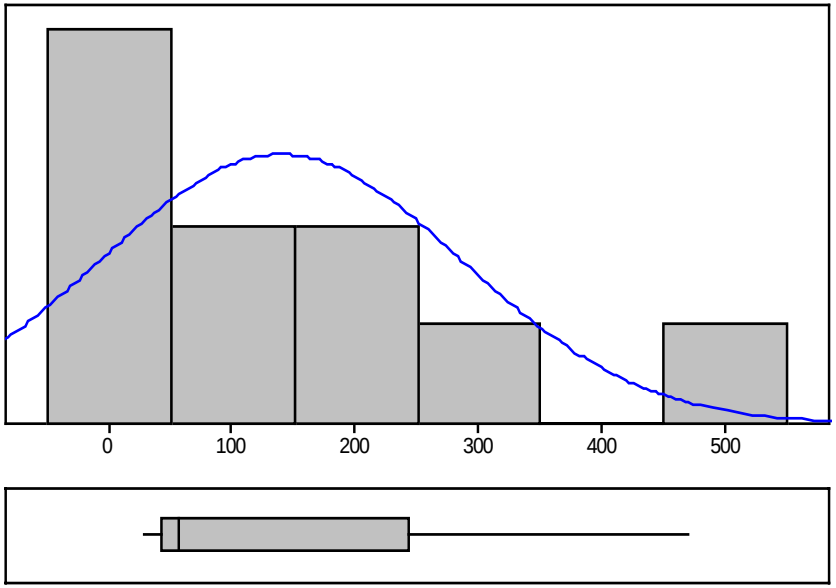
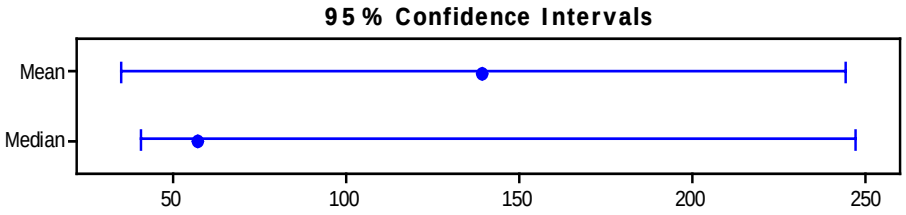


Figure 5-11. Logarithms of data set as a probability plot.

Here's what the Forest View BOD Data looks like (Provided by Tony Stahl, KDHE):

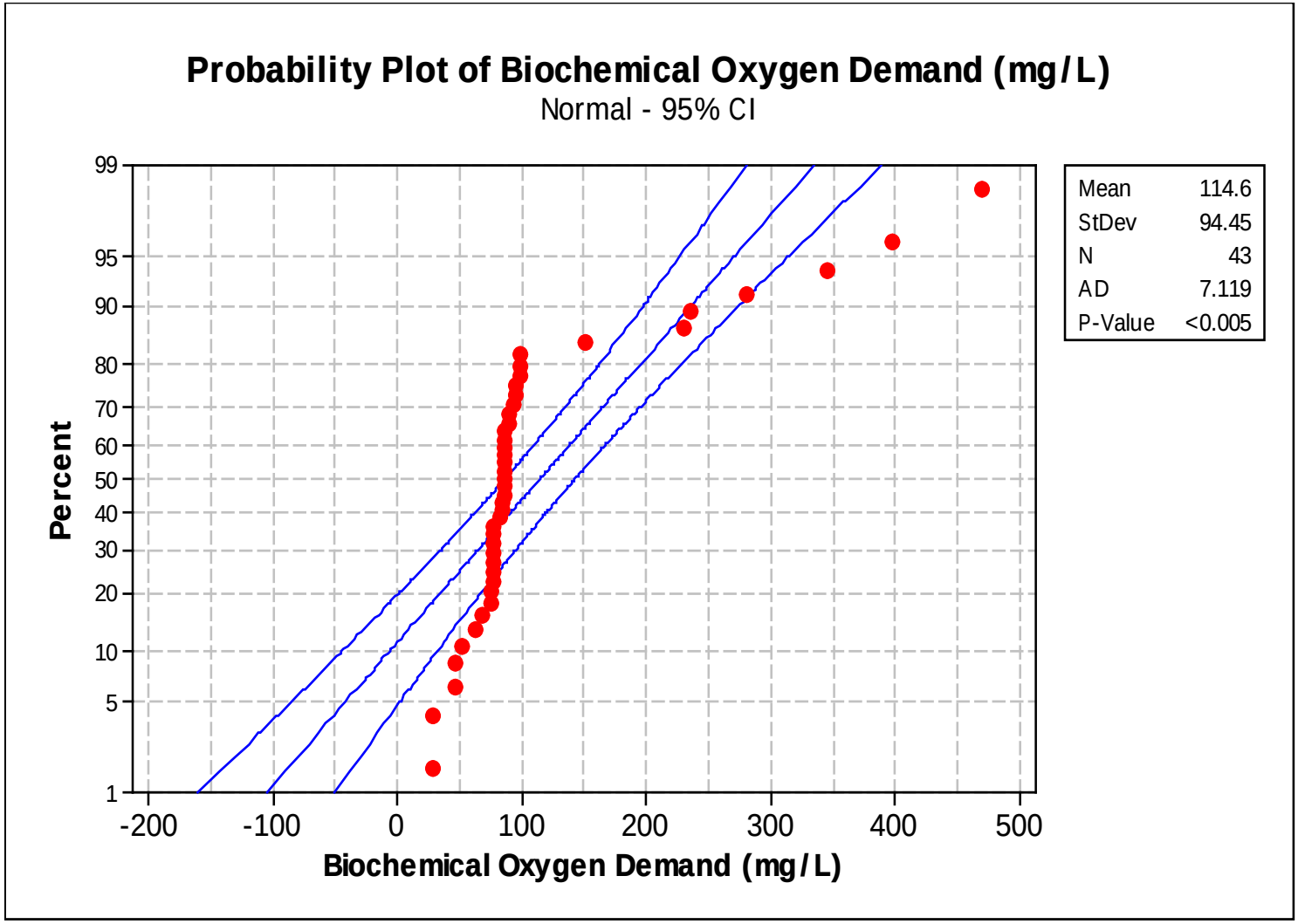
Summary for Biochemical Oxygen Demand (mg/L)





Anderson-Darling Normality Test	
A-Squared	0.93
P-Value	0.011
Mean	139.50
StDev	146.32
Variance	21409.83
Skewness	1.51461
Kurtosis	1.81854
N	10
Minimum	28.00
1st Quartile	42.50
Median	57.00
3rd Quartile	242.50
Maximum	470.00
95% Confidence Interval for Mean	
34.83	244.17
95% Confidence Interval for Median	
40.84	247.12
95% Confidence Interval for StDev	
100.64	267.13

Here's the BOD probability plot (Provided by Tony Stahl, KDHE):



Conclusion: **Must use non-parametric statistics to evaluate data. There are different tests to use (see selection guidelines in the list of references given in guidance paper which accompanies this presentation on the BWM website).**

No. 4 – Description and demonstration of the Kendall's tau method; but first, here are some of its features (as per ITRC Guidance and mine):

1. It is a **non-parametric** test for **monotonic** trends (i.e, they are either increasing or decreasing with time, but are not cyclic).
2. A minimum of 8 to 10 data points is recommended.
3. Results not impacted by magnitude of **extreme values or outliers**; and, not impacted by **censored (non-detects**) data.
4. Strength of trend (or lack of) can be determined.
5. An alternate test (**Kendall's tau b**) can be used for data sets where duplicate values (ties) exist.
6. Can be compared to other non-parametric statistics, e.g., **Thiel-Sen, Pearson's r** or **Spearman's rho**.

Here's the Kendall's tau formula:

$\tau = S/n(n-1)/2 = 2(S)/n(n-1)$ where:

$S = P - M$ = test statistic

n = number of data points

P = **increasing** numbers designated: +'s

M = **decreasing** numbers designated: –'s

$n(n-1)/2$ = number of possible comparisons or pairs.

Let's apply the formula to Forest View BOD data.

Revised Figure 1 - Forest View Time Series BOD Plot

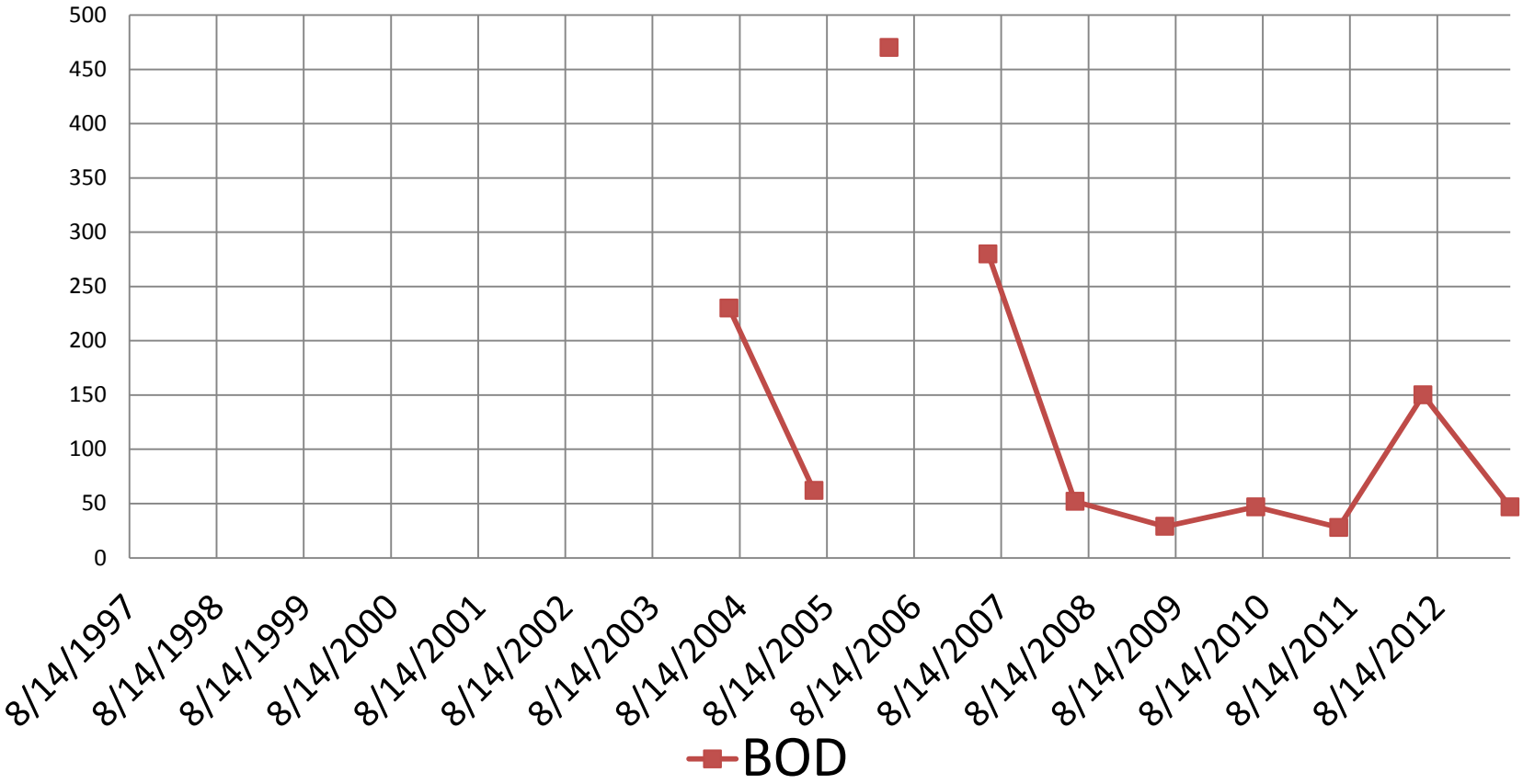


Table 1 - Using the last 7 data points in the previous figure (although 8 to 10 is the suggested minimum), the increasing and decreasing values are given + or – signs, respectively (same values are designated as a "tie" and are given a 0 sign); therefore,

T = 2(7 – 13) /7(7 – 1) = 2 (- 6)/7(6) = - 12/42 = - 0.286

Note that strong linear correlations of 0.9 or above correspond to tau values of about 0.7 or above (5); hence the calculated value of – 0.286 does not indicate a strong trend.

X (Date)	Y (mg/L)	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	Sum
2007	280							
2008	52	-						
2009	29	-	-					
2010	47	-	-	+				
2011	28	-	-	-	-			
2012	150	-	+	+	+	+		
2013	47	-	-	+	0	+	-	
P		0	1	3	1	2	0	7
M		6	4	1	1	0	1	13

Conclusions from Kendall's tau calculation:

1. It can be seen that if the BOD values decreased steadily, where the next value is lower than the preceding value (and whether the decreases are linear or non-linear), the P sum would be 0 and M would be – 21. This would give a Kendall's tau of – 1.00 indicating a perfect negative trend and vice versa if the values were increasing, or T = + 1.00. Values close to zero indicate no trend or equilibrium as far as PCC considerations are concerned.

Conclusions from Kendall's tau calculation (continued):

2. The strength of the tau calculation is reflected in the magnitude of S, i.e. as it approaches the possible number of comparisons or pairs, n(n-1)/2, the value of T approaches – 1.00 or + 1.00.
3. Ties reduce S and hence the values S and T.
4. The more data, with a proportional number of ties, gives No. 3 results; however, with less ties, the strength of S and T increases.

Table 2, which follows, is a summary of various statistics generated by Minitab © (as used by Tony Stahl) for the actual phase data shown below along with hypothetical values.

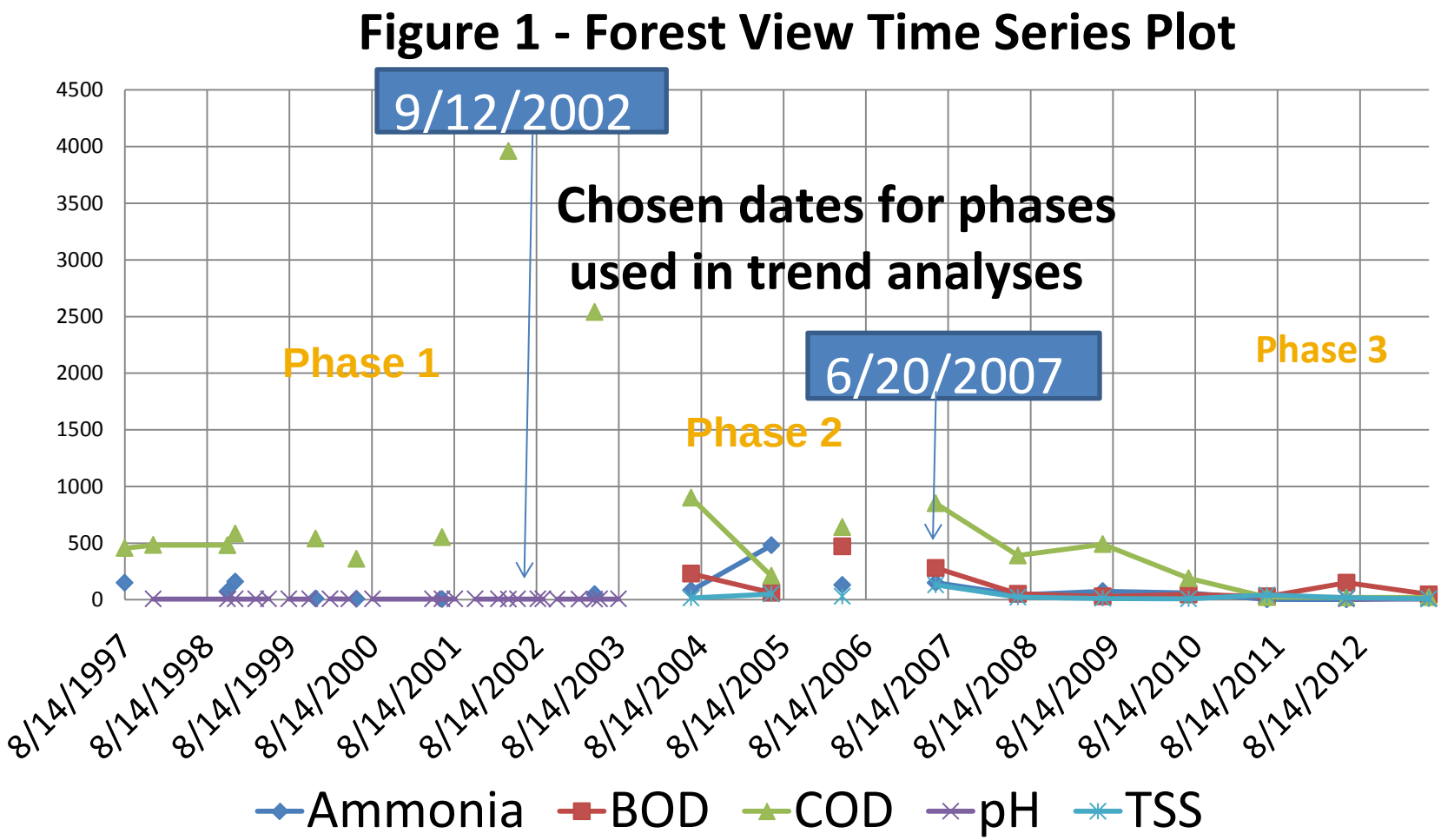


Table 2 – Minitab © results for Forest View

Phase/ Test	Start Date	End Date	N	Statistics			
Phase 1	**8/14/1997**	**8/21/2002**	**23**	*p* value	Slope	***tau***	Trend
Ammonia				**0.008**	**-16.818**	**-0.500**	**decreasing**
BOD				**0.851**	**0.000**	**-0.050**	**no trend**
COD				**0.010**	**26.294**	**0.483**	**increasing**
pH				**0.001**	**-0.109**	**-0.617**	**decreasing**
TSS				**0.014**	**5.284**	**0.467**	**increasing**
Phase 2	**9/12/2002**	**5/2/2006**	**13**				
Ammonia				**0.000**	**30.793**	**0.872**	**increasing**
BOD				**0.326**	**34.783**	**0.218**	**no trend**
COD				**0.902**	**13.062**	**0.038**	**no trend**
pH				**0.124**	**0.166**	**0.333**	**no trend**
TSS				**0.001**	**-13.000**	**-0.705**	**decreasing**
Phase 3	**6/20/2007**	**6/15/2013**	**7**				
Ammonia				**0.072**	**-15.525**	**-0.619**	**no trend**
BOD				**0.448**	**-2.413**	**-0.286**	**no trend**
COD				**0.017**	**-121.616**	**-0.810**	**decreasing**
pH				**0.031**	**-0.097**	**-0.714**	**decreasing**
TSS				**0.172**	**-7.238**	**-0.476**	**no trend**

Comments on Table 2 Results

1. **Hypothetical data** were added to the existing Forest View data base for all the parameters in order to demonstrate the completion of the statistical requirements for BWM PCC trend analysis. **Actual data** need to be added for future trend analysis.
2. **Phase 2** includes only 4 years of data which is one year less than the required 5 year phase requirements for reduction &/or termination of PCC activities.
3. Only BOD shows "**no trend**" for all phases which means it satisfies the BWM criterion for PCC termination.
4. All leachate parameters must have 3 consecutive 5 year "**no trends**" to satisfy PCC termination requirements; **and** some form of treatment (e.g., with ammonia) **may** be needed to insure future protection of human health & environment.

No. 5 – Statistical Resources include:

1. USGS, EPA and ITRC guidance documents.
2. Non-parametric statistics textbooks by Helsel and by Conover.
3. On-line sources on basic statistics.
4. Statistics software, e.g., Minitab ©.
5. BWM website policy & guidance documents.

PS: You need good data to get good statistics! Start now!

Questions

You may e-mail or call me if you have later questions and/or comments.

Kansas
Department of Health
and Environment

Contact information:

Dr. Carl E. Burkhead, 785-296-6898,
cburkhead@kdheks.gov

Our Mission: To protect and improve the health and environment of all Kansans.