PNDEP CEM Work Shop: Tools for Showing Plume Stability/Trends

"It is far better to have an approximate answer to the right question than a precise answer to the wrong question..." — John Hauser

1. VISUAL ANALYSIS

<u>Overview</u>

Using industry standard data analysis tools (i.e., Excell), plot data (groundwater level, concentration) vs time. This allows for visual identification of trends over time. These types of plots are commonly referred to as time-series plots.

Advantages

- ➢ Easily created.
- Minimal data requirements (can be as little as two data points, although more are preferred).
- Can show different components of project (VE system turned on, VE system turned off, etc.) to aid in evaluating trends.

Disadvantages

- > Data outliers can inadvertently influence "trend."
- > Does not account for variability in seasonal groundwater fluctuations.
- Trends identified can be







2. LINEAR REGRESSION

<u>Overview</u>

A parametric statistical technique used to estimate a trend via a linear relationship between multiple data points (sample analytical results). A line with a positive slope indicates an increasing trend, whereas a negative slope is indicative of a decreasing trend. Assumptions of linear regression are as follows:

- 1. The difference between each concentration measurement and its predicted value from the regression equation (residuals) are approximately normal in distribution.
- 2. Missing data and ND's are not part of data set.

The following conditions should be met prior to concluding a resulting trend:

- 1. The residuals (R) are approximately normal/reasonably symmetric in distribution.
- 2. A scatter plot of residuals vs concentrations yields a scatter cloud of generally uniform thickness.
- 3. A scatter plot of residuals vs time yields a scatter cloud of generally uniform thickness.
- 4. A minimum of eight measurements.

If the above conditions are met, the following can generally be applied with respect to trend interpretation:

- Where y=mx+b represents the trend line, and m=slope, if m is negative, the trend (slope) is decreasing.
- > Conversely, if m is positive, the trend is increasing.
- A higher m value (steeper slope) indicates a more rapid rate of degradation/contamination.
- When the above items cannot be met in an approximate sense, a non-parametric trend method should be utilized (Mann-Kendall).
- A smaller R² value indicates a less accurate trend line (if R² = 0, then the trend line would have no linear relationship). The larger the R² value, the less the amount of variation/deviation in the dataset from the trend line, and the more reliable the trend line is (If R²=1, the dataset is defined as linear).

<u>Advantages</u>

- Least squares regression is the most commonly used regression method. Calculates a best fit line for the observed data by minimizing the sum of the squares of vertical deviations from each data point to the line.
- Can be used site-wide or for individual wells, but is best suited for individual well analysis.
- > Principles can be applied to site-wide plume characteristics for site wide analysis
- > Relatively simple trend analysis/data requirements.

Disadvantages

- Data outliers can inadvertently influence "trend." Data input/output must be QA/QC'd prior to determining accuracy of trend (i.e., checking the residuals and/or visual).
- Extrapolation of data is not recommended in support of project decision making, yet can be useful as a general forecasting tool.
- > Trends are dependent on data quality and user interpretation.
- > Does not account for ND or missing data.

Online/Free-ware availability

Regression calculators and/or freeware are readily available online. Much of the freeware discussed in the following slides are equipped with regression analysis tools. It is also easily set up using Microsoft's Excel program.





Sources

Peter J. Brockwell, Richard A. Davis , *Time Series: Theory and Methods, Second Edition* Springer, 2009

Environmental Protection Agency, Office of Resource Conservation and Recovery, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities – Unified Guidance*, March, 2009

http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm

3. MANN-KENDALL (M-K)

<u>Overview</u>

M-K analysis is a non-parametric test for identifying trends in time-series data. In short, the analysis compares relative magnitudes of sample data (not the data values themselves). If an increasing trend exists, the sample taken first from any randomly selected pair of measurements should on average, have a lower concentration than the measurement collected at a later time.

The M-K statistic (S) is given by examining all possible pairs of data points (concentrations) and scoring each pair by assigning a value (identical values = 0, earlier value > later value = -1, earlier value < later value = 1). Summing the total of the assigned values gives S. A positive S suggests an upward trend, while a negative S suggests a negative trend. The larger the value of S (+/-), the stronger the level of confidence that the trend is legitimate.

Assumptions are as follows:

- 1. Only relative magnitudes are required (not actual concentrations) to rank the data.
- 2. ND should be treated as a common value lower than any detected values.
- 3. At least four data points must be analyzed.

Advantages

- > Relatively simple data requirements.
- ➢ User friendly.
- > Quantifies confidence level in trends based on data.
- > Enables quick identification of trends, historic and/or recent.

Disadvantages

- Does not account for site specific characteristics such as seepage velocity or well location.
- > Analyzes only a single data point (monitoring well).
- Seasonal effects are not accounted for which can incorrectly influence trend (groundwater fluctuation through "smear zone").
- Must address ND values...they need to be the same.

Online/Free-ware availability

Mann-Kendall free-ware is available online. Two good options are:

- 1. GSI Mann-Kendall Toolkit http://www.gsi-net.com/en/software/free-software/gsi-mann-kendall-toolkit.html
- 2. Washington State Department of Ecology (Package A) http://www.ecy.wa.gov/programs/tcp/policies/pol_main.html

Mann-Kendall analysis tools are also provided in the freeware packages discussed in the following slides.

Example Output

1. <u>(GSI)</u>



2. <u>(WSDE)</u>

Г

Module1: Mann-Kendall Trend Test for Plume Stability (Non-parametric Statistical Test)

Site Name:	ABC Corporation
Site Address:	Reno, NV
Additional	
Description:	

Well 2

Well (Sampling) Location?

Level of Confidence (Decision Criteria)? 85%

Monitoring Well Information: Contaminant Concentration at a well: Quarterly sampling recommended.

	Clear all dates	Hazardous Substances (unit is ug/L)								
Sampling Event	Date Sampled	Benzene	MTBE							
#1	2-Apr-10									
#2	8-Jul-10	44	110							
#3	20-Oct-10	73	80							
#4	20-Jan-11	19	18							
#5	18-Apr-11	64	11							
#6	24-Jul-11	9.10	5							
#7	18-Oct-11	8.90	5							
#8	17-Feb-12	18	3							
#9	7-May-12	51	3							
#10	16-Jul-12	35	2							
#11	22-Oct-12	11	2							
#12	23-Jan-13	21	1							
#13	11-Apr-13	20	2							
#14	10-Jul-13	20	2							
#15	4-Nov-13	23	2							
#16	23-Jan-14	8.0	1							

Mann-Kendall Non-parametric Statistical Test Results

Hazardous Substance?	Benzene	MTBE				
Confidence Level Calculated?	88.00%	100.00%	NA	NA	NA	
Plume Stability?	Shrinking	Shrinking	NA	NA	NA	
Coefficient of Variation?			n<4	n<4	n<4	
Mann-Kendall Statistic "S" value?	-26	-86	0	0	0	
Number of Sampling Rounds?	15	15	0	0	0	
Average Concentration?	28.33	16.36	NA	NA	NA	
Standard Deviation?	20.61	32.75	NA	NA	NA	
Coefficient of Variation?	0.73	2.00	NA	NA	NA	
Blank if No Errors found			n<4	n<4	n<4	



<u>Sources</u>

Environmental Protection Agency, Office of Resource Conservation and Recovery, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities – Unified Guidance*, March, 2009

GSI Environmental, Inc., GSI Mann-Kendall Toolkit for Constituent Trend Analysis, User's Manual, Version 1.0, November, 2012

Washington State Department of Ecology, User's Manual: Natural Attenuation Analysis Tool Package for Petroleum-Contaminated Ground Water, July, 2005

4. WASHINGTON STATE DEPARTMENT OF ECOLOGY: NATURAL ATTENUATION ANALYSIS TOOL PACKAGE FOR PETROLEUM-CONTAMINATED GROUND WATER

Overview

Comprehensive site evaluation tool based on statistical methods applied to site-specific data that accounts for historical and current data as well as hydrogeologic factors (e.g., wells, seepage velocity), and potential receptors. Analyzes individual wells, plume with respect to temporal/spatial indicators and stability. Using statistical trend analysis, helps identify temporal trends, plume characteristics (shrinking/expanding), time until target concentrations are met, influence of groundwater, evaluation of geochemical indicators, and graphical presentation of historical groundwater data. The program is divided into two packages, A and B.

Package A (Modules 1, 2, and 3) analysis tool will conduct the following:

- > Non-parametric statistical tests for plume stability at each well.
 - Mann-Kendall test (previously discussed)
 - o Mann-Whitney U-test
- > Graphical presentation of historical ground water data.
 - Plot of temporal ground water analytical and elevation data vs. time to assess the plume status and the impact of ground water elevation fluctuation on contaminant concentrations at each well
 - Plot of spatial ground water analytical data vs. distance (for multiple wells) to assess the overall plume status
- > Evaluation of geochemical indicators.
 - o Estimate of expressed assimilative capacity at multiple wells
 - Simultaneous plot of concentrations of contaminant and geochemical indicators vs. distance (at multiple wells) to demonstrate biodegradation clearly
- > Temporal trend (regression) analysis at each well.
 - Estimate of an average and a range of (kpoint) point decay rate (1st-order) constant for both the best-fit and a given one-tailed confidence level at each well
 - Temporal prediction at each well location under a given confidence level
 - Estimate of an average and a range (under a given confidence level) of restoration time to reach the cleanup goal at each well
 - Calculation of the correlation coefficient and confidence level (with the Pearson's correlation coefficient) of log-linear regression analysis (for a plot of concentration vs. time at each well)



Note: Modules are not linked each other.



7/2005: Version 1.0; User's Manual: Natural Attenuation Analysis Tool Package for Petroleum-Contaminated Ground Water

Package B (Modules 4, 5, and 6) analysis tool will conduct the following calculations:

- Estimate of source mass from sampling data: for unsaturated, smear, and dissolved zones.
- Under 1-D (transformed from 2-D): steady state/continuous source assumption for only stable.
 - plume (with Buscheck and Alcantar model: see footnote on page 33 of this User's Manual)
 - Plot of the concentration vs. distance
 - Estimate of an average and a range of (λ) biodegradation rate constant
 - Estimate of an average and a range of (k) bulk attenuation rate (1st-order) constant under steady state (stable plume)
 - Estimate of a percent mass removal rate by biodegradation alone
 - Temporal and spatial prediction as a function of time and well location
 - Estimate of a target source concentration in order to reach a target level at a receptor location under given restoration time
- Under 2-D; transient state (with modified Domenico model) for shrinking and stable (or any type) plumes:
 - Estimate of a biodegradation rate constant (λ) by calibration via chi-square statistics for best-fit to the normalized concentration of consecutive multiple wells by 1st-order decay model
 - Estimate of a percent mass removal rate by biodegradation alone with 1st-order decay model and instantaneous reaction model (via the calculation of mass flux)
 - Estimate of a temporal/spatial prediction at a receptor location by 1st-order decay model and instantaneous reaction model
 - Estimate of a plume stabilization time (half time to reach the steady state) at a receptor location
 - Estimate of a restoration time to reach a target level at a receptor location by 1storder decay model and instantaneous reaction model
 - Estimate of a target source mass amount (amount of mass that should be removed from the current source zone) in order to reach a target level at a receptor location under a given restoration time by 1st-order decay model and instantaneous reaction model
 - Estimate of a contaminant mass loading rate (as a function of x-distance and time) to the adjacent surface water body by 1st-order decay model





7/2005: Version 1.0; User's Manual: Natural Attenuation Analysis Tool Package for Petroleum-Contaminated Ground Water

Advantages

- ➢ User friendly
- > Quantifies confidence level in trends based on data
- > Enables quick identification of trends, historic and/or recent
- Comprehensive site analysis
- > Incorporates hydrogeologic site data and potential receptors in analysis
- > Geochemical/Biodegradation assessment/modeling capability
- 2-D Modeling capability

Disadvantages

- Complex data requirements (depending on module)
- Time intensive (depending on module)
- > Does not use real world coordinates in 2-D modeling components
- > Does not provide site optimization recommendations

Online/Freeware availability

Free-ware is available online:

http://www.ecy.wa.gov/programs/tcp/policies/pol_main.html



Module 6	dule 6: Calculation of Mass Removal Rate by Biodegradation							Site Name: Dummy XYZ site					
Main	Print View	End	End							1234, Olym Powzowa	pia, WA 985	501	
Transvers Distance, y	Transverse Ground water Concentrations in Plu Distance x direction (ft) Distance from Source x direction							Hazardous Substance: Benzene Ime (ug/L at z=0)				Choose a Model	
Jistanec,	0.0	55.0	110.0	165.0	220.0	275.0	330.0	385.0	440.0	495.0	550.0	No De	aredation
100.0	36.3	39.6	48.8	37.2	23.9	14.1	8.0	4.3	2.3	1.2	0.6	Mo Degradation Model	
50.0	4669.0	1681.7	789.4	372.3	176.8	84.3	40.4	19.4	9.3	4.5	2.1	1 at Or	dar Dagay
0.0	4669.0	2272.0	1103.6	533.2	256.3	122.8	58.7	28.1	13.4	6.4	3.0	TSI ON	ler Decay
-50.0	4669.0	1681.7	789.4	372.3	176.8	84.3	40.4	19.4	9.3	4.5	2.1		louel
-100.0	36.3	39.6	48.8	37.2	23.9	14.1	8.0	4.3	2.3	1.2	0.6	Insta	ntaneous
MASS	3.00E+4	7.56E+3	3.68E+3	1.79E+3	8.71E+2	4.23E+2	2.06E+2	9.99E+1	4.85E+1	2.35E+1	1.14E+1	React	ion Model
FLUX			T	arget Ground wa	ater Conc, ug/L :	5.0	, ←L →	Enter Sim	ulation Time, yr:	12	,	1	
(mg/day)				- Modeled Area	a Length (L), ft :	550.0		*Displayed	Model is 1s	t-order Dec	av.		
				Modeled Area	a Width (W) ft	200.0							
5000 4000 4000 50000 5000 5000 5000 5000 5000 5000 5000 5000						= Plur Oxygen na	Plume Ma me Mass Rem Change in E Nitrate na Contam. Mass i Current Volu Flowrate of V	- Actual 1 - Actual 1 noved by Bio 	degradation: Plume Mass: degradation: eptor/Byprodu sutrate na ce (t=0 years): v (t=12 years): vater in Plume: a Source Zone:	96.2 5.3 90.9 uct Mass, kg Methane na 200.0 103.8 10.4 1.960	kg kg (94 %) Manganese na kg kg ac-ft ac-ft		
Plot Dat	a > Target				550								





Location of Receptor #1 is 50ft x-direction & 10ft y-direction; Location of Receptor #2 is 200ft x-direction & 20ft y-direction



 4. Temporal/Spatial Prediction along Plume Centerline (@ y & z = 0) with 1st-order Decay Model

 Modeled Overall Plume Centerline Distance to evaluate, ft

 Modeled Overall Simulation Time to evaluate, year

 12



Sources

Washington State Department of Ecology, User's Manual: Natural Attenuation Analysis Tool Package for Petroleum-Contaminated Ground Water, July, 2005

5. MONITORING AND REMEDIATION OPTIMIZATION SYSTEM (MAROS)

Overview

Developed by GSI for the Air Force Center for Environmental Excellence (AFCEE), in accordance with the AFCEE Long Term Monitoring Optimization Guide.

Comprehensive site evaluation tool based on statistical methods applied to site-specific data that accounts for historical and current data as well as hydrogeologic factors (e.g., wells, seepage velocity), and potential receptors. Analyzes individual wells, plume with respect to spatial indicators and stability, and site optimization. Using database trend analysis, helps identify constituents of concern (CoC), significance of temporal trends, redundancy of data points (monitoring wells), adequate sampling frequency, and data gaps.

MAROS will conduct the following calculations:

- Summary Statistics for Individual Wells
 - Calculates the detection frequency, date range of data, maximum concentration, range of concentrations, and date of maximum result for up to 5 COCs for all wells
 - Summary Statistics using Kaplan-Meier Method: Mean, median standard deviation and percentiles for individual well data are calculated using the Kaplan-Meier method to account for datasets with a higher percentage of non-detect (ND) data
 - Outliers for Individual Wells: uses Dixon's method to identify high or low outliers in a dataset
 - Data Distribution: MAROS Uses the Shapiro-Wilk method to identify individual well datasets that do not have Normal or Log-normal data distributions
- > Trend Analysis for Individual Wells
 - Mann Kendall
 - o Linear Regression
- > Data Sufficiency for Individual Wells
 - Cleanup Status: Sequential T-Test and Student's T-Test are used to determine if concentrations are statistically below the cleanup goal
 - Power Analysis: Estimates how many more samples may be required to demonstrate location is statistically below the cleanup level
 - Prioritizes well importance using a qualitative method based on individual well statistics and well monitoring objectives
- Moment Analysis

- Uses the full dataset to estimate moments and the Mann-Kendall trends of each metric.
- Total dissolved mass in the plume
- Center of mass: Coordinates of the center of mass
- Spread of mass about the center of the plume
- Evaluation of aggregate concentration trends for source area, tail and User-defined well groups
- Spatial/Temporal Optimization
 - Provides several qualitative and quantitative metrics for identifying redundant monitoring locations and for identifying areas of high uncertainty that may require more monitoring locations
 - Provides several qualitative and quantitative metrics for assessing appropriate sampling frequency for well networks



Advantages

- Comprehensive site analysis
- > Incorporates hydrogeologic site data and potential receptors in analysis
- Recommendations for sampling optimization (potential for cost reduction)
- Recommendations for potential well locations (data gaps)
- > EPA accepted data analysis tool for project decision support
- Uses real world coordinates in 2-D modeling components

Disadvantages

- Complex data requirements
- > Time intensive
- > Does not analyze geochemical/biodegradation components (work-around required)

Online/Freeware availability

http://www.gsi-net.com/en/software/free-software/maros.html

🔡 Moni	itoring and Remediation Optimization System (MARDS)	
	Linear Regression Plot	
	Select a well and chemical below to graph. The concentration trend result in the box below reflects the chemical and well chosen to be graphed.	
	Select: Well MW-11 Chemical BENZENE	<u> </u>
	Date 30 30 30 30 30 30 30 30 30 30 40 30 30 40 40 40 40	Graph Type C Log O Linear
	8.0E+00 7.0E+00 6.0E+00 5.0E+00 4.0E+00 3.0E+00	Graph View Data
	5 2.UE+00 - 5 1.0E+00 - 5 1.0E	5 95E-04
NLYSIS	0.0E+00 J	Confidence in Trend: 98.9%
D ANA	Linear Regression Trend:	COV:
RENI	Note: Increasing (); Probably Increasing (P); Stable (S); Probably Decreasing (PD); Decreasing (D); No Trend (NT); NotApplicable (N/P) - due to institute it data.	
F	<< Back Next >> View Report	<u>H</u> elp







Sources

GSI Environmental, Inc., Air Force Center for Engineering and the Environment, Monitoring and Remediations System (MAROS), User's Guide and Technical Manual, September, 2012